

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: October 14, 2003, 15:48:26 ; Search time 23109 Seconds
(without alignments)
10933.318 Million cell updates/sec
Title: US-09-754-997A-1
Perfect score: 6176
Sequence: 1 atggcgccggcgacacggg.....taataaactattttccagac 6176
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2888711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rod.*
- 36: em.htg_man.*
- 37: em.htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	6176	100.0	6176	6	AX191259	AX191259 Sequence
2	6176	100.0	6176	10	AF176694	AF176694 Mus muscu
3	6038.8	97.8	6219	10	AB052620	AB052620 Mus muscu
4	6025.8	97.6	6222	10	AB052621	AB052621 Mus muscu
5	5810.2	94.1	6301	10	AK122535	AK122535 Mus muscu
6	3756	60.8	3756	6	AX191303	AX191303 Sequence
7	3396.6	55.0	6485	9	AB052622	AB052622 Homo sapi
8	2796	45.3	2796	6	AX191261	AX191261 Sequence
C 9	2743.2	44.4	115135	10	AC112161	AC112161 Mus Muscu
C 10	2743.2	44.4	158055	2	AC110235	AC110235 Mus Muscu
11	2691	43.6	5549	9	AB046848	AB046848 Homo sapi
12	2681.8	43.4	3741	6	AX18406	AX18406 Sequence
C 13	1843.2	29.8	224227	2	AC106509	AC106509 Rattus no
C 14	854.8	13.8	176282	2	AC068507	AC068507 Homo sapi
15	854.8	13.8	198295	9	AC105129	AC105129 Homo sapi
16	825	13.4	825	6	AX191263	AX191263 Sequence
C 17	659.4	10.7	72149	2	AC122776	AC122776 Mus muscu
18	632.2	10.2	2560	9	BC040648	BC040648 Homo sapi
19	571.2	9.2	72149	2	AC122776	AC122776 Mus muscu
C 20	415	6.7	116580	10	AC112162	AC112162 Mus Muscu
21	378.8	6.1	216302	2	AC107847	AC107847 Mus muscu
22	317	5.1	2967	9	BC042054	BC042054 Homo sapi
23	316.6	5.1	2976	6	AR237558	AR237558 Sequence
24	316.6	5.1	3210	6	AR237553	AR237553 Sequence
25	316.6	5.1	3219	6	AR237561	AR237561 Sequence
26	316.6	5.1	3874	6	AR237568	AR237568 Sequence
27	316.4	5.1	2700	6	AX497857	AX497857 Sequence
28	315.4	5.1	3453	6	AR237556	AR237556 Sequence
29	308.4	5.0	3198	10	BC053057	BC053057 Mus muscu
30	306.8	5.0	3146	10	AF026465	AF026465 Mus muscu
31	288	4.7	288	6	AX191277	AX191277 Sequence
32	287.4	4.7	2729	9	AK095529	AK095529 Homo sapi
33	252	4.1	252	6	AX191281	AX191281 Sequence
34	249	4.0	249	6	AX191273	AX191273 Sequence
35	249	4.0	249	6	AX191275	AX191275 Sequence
36	246	4.0	246	6	AX191279	AX191279 Sequence
37	243	3.9	243	6	AX191265	AX191265 Sequence
38	217	3.5	2481	6	AR237560	AR237560 Sequence
39	217	3.5	2715	6	AR237555	AR237555 Sequence
40	217	3.5	2724	6	AR237562	AR237562 Sequence
41	217	3.5	2958	6	AR237557	AR237557 Sequence
C 42	215.8	3.5	147760	9	AC011846	AC011846 Homo sapi
43	215.8	3.5	176282	2	AC068507	AC068507 Homo sapi
44	213.8	3.5	2139	6	AR237563	AR237563 Sequence
45	213.8	3.5	2382	6	AR237566	AR237566 Sequence

ALIGNMENTS

RESULT 1
AX191259

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

AX191259
Sequence 1 from Patent WO0149714.
AX191259.1
GI:15209510

Mus musculus (house mouse)
Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL

1
Salbaum,J.M.
Nope polypeptides, encoding nucleic acids and methods of use
Patent: WO 0149714-A 1 12-JUL-2001;

linear PAT 15-AUG-2001

Neurosciences Research Foundation Inc. (US)

FEATURES
Source

Location/Qualifiers

1..6176

/organism="Mus musculus"

/mol_type="Genomic DNA"

/db_xref="taxon:10090"

1..3759

CDS

/note="unnamed protein product"

/codon_start=1

/protein_id="CAC51120.1"

/db_xref="GI:15209511"

/translation="NARADTGRGLLVLFCLLSARGELPLPQETTVKLSCDEGPLQVI
LGEQAVLDCTAGATAAGPPTVTMSKGDVLEHENHLLPNSLSPLEQSDS
DDEALRTVKTVEGYSCLASPLGASQVAVKLAILEDPLSPESQIVBENGTA
RECHTKGLPAPLITWEDQVTPPEPRILITLPMKLLQILDVQSDAGSYRCVATNSA
QRFSQESLITVALRSLSEATRGQDVVIAAPENTTVSGQNVMECVASAPTPFVSW
VRQDKPSTIDVIGRNTLLIASQPHSGVYVCRANKPLTRDPATAAEULALAP
AISOPEALSSTRASTARFVCRASGEPRPALHWDGIPLRPNRGVKGQGGGSLVIT
QIGLQADYQCCASNSAGTACAAAPLAVVREGILPSAPTRVATPLSSSVLWAMER
PELHSEQIIGFSLHYOKARGVDNVQYQFVANNDDTELQVRLDLEPNTDYEFVYVAYSL
CASRTSSPALVHTLDDVPSAAPOLTLSSPNPSDIRVAMLPSSLSNGQVLKYIEYG
LKEQDVSTEVPGNETQITLNSLPKNKYRVRISAGTAGYGVPSQWQHRTPOVHN
QSHVFPABULKRAKMSLVVSWQPPPHPTQISGYKLYMGVEVTEESADGDRPPGR
GDQAVDQVPLKKKQVQYELTQVPRPYEVKLVAFKNHEDGYAAVWKGTETKAPTP
DLPIQRGGLPPLPAHVAESNSTSIWLRKKPDPFTTVKIVNYTVRFPGWGLRNASLPT
YVTSGEDILIGLKPFTKYFAVQSGVMDMDPFGSVVERSTLDPDRSTPSSDLRLS
PLTPSTVRLWCPTPEPGEIVELYLISNNHTOPEHOWTILLTTGENIFSAREVHGLS
DTRYFFKMGATEVPGPSPRLQDVITLQETPSDLDVHVTGIIIVGLVGLLCLLAC
MCAGLRQSHREALPGLSSGTPGNPALYTRARLGPSPVPAHHELESILPRPQDMS
PPSDEVEDAEHSLMGGSVSDCRGHSKRKI SWAQAGPNWAGSCLSPQSGRPPA
LTRALLPAGTQOTLLOALVYDGI KSNRGKKPSACRNOVEAEVIVHSDFGASKGCP
DLHLQDLPEPELTAETLPSTSGAVDLSQGDMLRGELGGCQPTTSGPERLTLCLPEAA
SNSCSDLOPSTALEAPGSKQKALCLPLTVSPSLPRAPVSSAQVP"

BASE COUNT 1316 a 1756 c 1634 g 1470 t

ORIGIN

Query Match 100.0%; Score 6176; DB 6; Length 6176;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6176; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGCGGGCGGACACGCGGCGCGGCTCTGGTGCTGACCTTCTGCTGCTGCTGCGG 60
DB 1 ATGGCGCGGGCGGACACGCGGCGCGGCTCTGGTGCTGACCTTCTGCTGCTGCTGCGG 60

QY 61 CGCGGGGAGCTGCCATTCGCCACGAGACAACTGTCTCAAGCTGAGCTGTGATGAGGACCC 120
DB 61 CGCGGGGAGCTGCCATTCGCCACGAGACAACTGTCTCAAGCTGAGCTGTGATGAGGACCC 120

QY 121 CTGCAAGTGATCTCGGGCCCTGAGCAGGCTGTGGTGCTGCACTGTGCGGCTACA 180
DB 121 CTGCAAGTGATCTCGGGCCCTGAGCAGGCTGTGGTGCTGCACTGTGCGGCTACA 180

QY 181 GCTGCTGGGCTCCGACCAAGGTTGATGAGCAAGGATGGAGACACTGTACTAGAGCAT 240
DB 181 GCTGCTGGGCTCCGACCAAGGTTGATGAGCAAGGATGGAGACACTGTACTAGAGCAT 240

QY 241 GAGAACCTGCACTGCTACCAATGGCTCCCTGTGGCTGTCTCACCCTCAGACAGNA 300
DB 241 GAGAACCTGCACTGCTACCAATGGCTCCCTGTGGCTGTCTCACCCTCAGACAGNA 300

QY 301 GACAGCGATGATGAGGAGCTCTTAGGATCTGGAAGTCACTGAGGGCAGCTATTCTGT 360
DB 301 GACAGCGATGATGAGGAGCTCTTAGGATCTGGAAGTCACTGAGGGCAGCTATTCTGT 360

QY 361 CTGGCCCAAGCGGCTAGGAGTGGTGGCCAGCAGGTTGTGTGTCAAGCTTGGCACA 420
DB 361 CTGGCCCAAGCGGCTAGGAGTGGTGGCCAGCAGGTTGTGTGTCAAGCTTGGCACA 420

QY 421 CTGGAAGACTTCTCTGACCCCGAGTCCAGATTGTTGGAGAGAGAACGGGACAGCACGC 480
DB 421 CTGGAAGACTTCTCTGACCCCGAGTCCAGATTGTTGGAGAGAGAACGGGACAGCACGC 480

QY 481 TTTGAATGACACACCAAGGGCTTCCAGCCCCCATCATTTCTTGGGAAAAGGACGAGTG 540

DB 481 TTTGAATGACACCAAGGGCTTCCAGCCCCCATCATTTCTTGGAAAAGGACCGAGTG 540

QY 541 ACCGTGCTGAGGAGCCCCGGCTCATCACTCTTCCCAAGTGGCTCTCCACATCTAGAT 600
DB 541 ACCGTGCTGAGGAGCCCCGGCTCATCACTCTTCCCAAGTGGCTCTCCACATCTAGAT 600

QY 601 GTCCAGGACAGTGATGAGGCTCTCTACCGCTGCGTGGCCACCAATTCAGCCCGCCCAACGA 660
DB 601 GTCCAGGACAGTGATGAGGCTCTCTACCGCTGCGTGGCCACCAATTCAGCCCGCCCAACGA 660

QY 661 TTACAGCAGGAGGCTCGCTCATCTGTGGCCCTCAGAGGGTCTTTTGGAGGCTACCAAGGG 720
DB 661 TTACAGCAGGAGGCTCGCTCATCTGTGGCCCTCAGAGGGTCTTTTGGAGGCTACCAAGGG 720

QY 721 CAGATGTGGTCATTTGTGGAGCCCGCAGAGAACCAACGAGTGTCTGGACAGATGTA 780
DB 721 CAGATGTGGTCATTTGTGGAGCCCGCAGAGAACCAACGAGTGTCTGGACAGATGTA 780

QY 781 GTGATGAGTGGTGGCTCTGTGACCCCAACCCCTTTTGTGTCTGGGTCCGACAGGAT 840
DB 781 GTGATGAGTGGTGGCTCTGTGACCCCAACCCCTTTTGTGTCTGGGTCCGACAGGAT 840

QY 841 GGAAGGCTATCTCCAGGATGTCTGCTGAGCCCGGCGGACCAATCTACTCATCGCCAGC 900
DB 841 GGAAGGCTATCTCCAGGATGTCTGCTGAGCCCGGCGGACCAATCTACTCATCGCCAGC 900

QY 901 CGCAGCTCTCGGCTCTGGAGTCTATGTCTGCGGAGGCAACAGCCCTCACGGTGCAC 960
DB 901 CGCAGCTCTCGGCTCTGGAGTCTATGTCTGCGGAGGCAACAGCCCTCACGGTGCAC 960

QY 961 TTGCGCACTGCGGCTGCTGAGCTCCGAGTGTCTGTCGCCCGAGCCATCTCCGAGCACCC 1020
DB 961 TTGCGCACTGCGGCTGCTGAGCTCCGAGTGTCTGTCGCCCGAGCCATCTCCGAGCACCC 1020

QY 1021 GAGGCGCTCTCGCGGACGCGGCGGCGGACGACCCGCGCTTCTGTGTGCGGCGCTCGGGAG 1080
DB 1021 GAGGCGCTCTCGCGGACGCGGCGGCGGACGACCCGCGCTTCTGTGTGCGGCGCTCGGGAG 1080

QY 1081 CCACGGCGCGCTGCTGAGTGTGCAAGCGGATCCGTTGCGACCCCAATGGGCGCTC 1140
DB 1081 CCACGGCGCGCTGCTGAGTGTGCAAGCGGATCCGTTGCGACCCCAATGGGCGCTC 1140

QY 1141 AAGTGTGAGGCGGCTGGCGGCTGCTGATCATCTCAAGTCTGCGCTTCAGAGACGCTGC 1200
DB 1141 AAGTGTGAGGCGGCTGGCGGCTGCTGATCATCTCAAGTCTGCGCTTCAGAGACGCTGC 1200

QY 1201 TACTACAGTGTGAGCAAGAAACAGCGCGGAACTGCTGTGCGCTGCGCCCTCGCG 1260
DB 1201 TACTACAGTGTGAGCAAGAAACAGCGCGGAACTGCTGTGCGCTGCGCCCTCGCG 1260

QY 1261 GTAGTGTGCGGAGGGGCTGCCAGCGCCCGGCTCGGGTTCAGAGCCAGCCGCTGAGC 1320
DB 1261 GTAGTGTGCGGAGGGGCTGCCAGCGCCCGGCTCGGGTTCAGAGCCAGCCGCTGAGC 1320

QY 1321 AGCTCTCTGTGTGCTGCTGCGAGCGGCTGAGTTGCAAGAGCAACATCAATTGC 1380
DB 1321 AGCTCTCTGTGTGCTGCTGCGAGCGGCTGAGTTGCAAGAGCAACATCAATTGC 1380

QY 1381 TTCTCTTCTACTACAAAAGCAAGGGAGTGGACAAATGTGGAGTACCAAGTTTACAGTA 1440
DB 1381 TTCTCTTCTACTACAAAAGCAAGGGAGTGGACAAATGTGGAGTACCAAGTTTACAGTA 1440

QY 1441 AACATGACACAGAGCTGAGGTTTCGGGACCTGGAAACCCAAACAGGATTTATGAGTTC 1500
DB 1441 AACATGACACAGAGCTGAGGTTTCGGGACCTGGAAACCCAAACAGGATTTATGAGTTC 1500

QY 1501 TACGTGTGGCTACTCCAGCTGGGGCCAGCGAACTCCAGCCAGCCCTGCTGAT 1560
DB 1501 TACGTGTGGCTACTCCAGCTGGGGCCAGCGAACTCCAGCCAGCCCTGCTGAT 1560

QY 1561 ACATGGAACGATGTCCCAGCGCAGCACCCAGCTTACCTTGTCCAGGCCCAACCCCTCG 1620

Db	1561	ACACTGGACGATGTCCCGACGGCAGCACCCGAGCTTACCTTGTCCAGCCCAACCCCTCG	1620
Qy	1621	GACATCAGGGTGGCATGGCTGCCCTGCCCTCCAGCTGAGCAATGACAGGTGCTGAAG	1680
Db	1621	GACATCAGGGTGGCATGGCTGCCCTGCCCTCCAGCTGAGCAATGACAGGTGCTGAAG	1680
Qy	1681	TACAAGATAGAGTACGGTTTGGGGAAGAAAGATCAGGTTTTCTCCACCGAGGTGCTTGA	1740
Db	1681	TACAAGATAGAGTACGGTTTGGGGAAGAAAGATCAGGTTTTCTCCACCGAGGTGCTTGA	1740
Qy	1741	AATGAGACACAACTTACGTTTAAACTCACTTCAGCCAAACAAAGTGTACCGAGTCCGGATT	1800
Db	1741	AATGAGACACAACTTACGTTTAAACTCACTTCAGCCAAACAAAGTGTACCGAGTCCGGATT	1800
Qy	1801	TCAGCTGGCACTGGCGCTGGCTATGAGAGTCCCTTCTCAGTGGATGACAGCACAGACCT	1860
Db	1801	TCAGCTGGCACTGGCGCTGGCTATGAGAGTCCCTTCTCAGTGGATGACAGCACAGACCT	1860
Qy	1861	GGTGTGCAACACAGAGCCATGTTCCCTTGGCCCTGCAGAAATTGAAGGTGAGGGCAAG	1920
Db	1861	GGTGTGCAACACAGAGCCATGTTCCCTTGGCCCTGCAGAAATTGAAGGTGAGGGCAAG	1920
Qy	1921	ATGGAGTCCCTGGTGTGATGACAGCGCCCTTCCAGCCACCCAGATCTCTGGATAC	1980
Db	1921	ATGGAGTCCCTGGTGTGATGACAGCGCCCTTCCAGCCACCCAGATCTCTGGATAC	1980
Qy	1981	AAACTCTACTGGGAGAGGTGGGAAACAGAGGAGGACAGATGGTACCCGCCCCCAGGG	2040
Db	1981	AAACTCTACTGGGAGAGGTGGGAAACAGAGGAGGACAGATGGTACCCGCCCCCAGGG	2040
Qy	2041	GGTGTGAGATCAAGCTTGGGACGTGCGGCCCGTGGCGGTGAAGAAAGTGAAGCAG	2100
Db	2041	GGTGTGAGATCAAGCTTGGGACGTGCGGCCCGTGGCGGTGAAGAAAGTGAAGCAG	2100
Qy	2101	TATGAATGACCCAGTTAGTCCCTGGGACAGCGCGTACGAGTGAAGCTCGTAGCTTCAAC	2160
Db	2101	TATGAATGACCCAGTTAGTCCCTGGGACAGCGCGTACGAGTGAAGCTCGTAGCTTCAAC	2160
Qy	2161	AAACAAGAGGACGGCTACGTGCTGTGAAGGGCAAGACGGAGAGGCGCCACGCCCA	2220
Db	2161	AAACAAGAGGACGGCTACGTGCTGTGAAGGGCAAGACGGAGAGGCGCCACGCCCA	2220
Qy	2221	GACCTGCTATCCAGAGGGGCCACCGCTGCTCCCTCCAGTGTCCACGAGAGTCAAAC	2280
Db	2221	GACCTGCTATCCAGAGGGGCCACCGCTGCTCCCTCCAGTGTCCACGAGAGTCAAAC	2280
Qy	2281	AGCTCCACTTCCATTTGGCTTGGTGAAGAGCAGACTTTACCACTGTCAAGATTGTC	2340
Db	2281	AGCTCCACTTCCATTTGGCTTGGTGAAGAGCAGACTTTACCACTGTCAAGATTGTC	2340
Qy	2341	AACCTACACTGTACGCTTCCGCCCCCTGGGGCTCAGGAATGCTTCCCTGGTCACTACTAT	2400
Db	2341	AACCTACACTGTACGCTTCCGCCCCCTGGGGCTCAGGAATGCTTCCCTGGTCACTACTAT	2400
Qy	2401	ACCAAGCTCTGGAGAACATTTCTCATTTGGGGCTGAAACCATTTTACCAAGTACGAGTTT	2460
Db	2401	ACCAAGCTCTGGAGAACATTTCTCATTTGGGGCTGAAACCATTTTACCAAGTACGAGTTT	2460
Qy	2461	GGGTACAGTCCCAAGAGTGATATGAGATGGGCCCCCTTGGCTCCGCTGTAGAACCGTCC	2520
Db	2461	GGGTACAGTCCCAAGAGTGATATGAGATGGGCCCCCTTGGCTCCGCTGTAGAACCGTCC	2520
Qy	2521	ACCTGGCTGACCGGCTTCAACACCTCTTCTGACCTGGCGCTGAGCCCTTGCACCA	2580
Db	2521	ACCTGGCTGACCGGCTTCAACACCTCTTCTGACCTGGCGCTGAGCCCTTGCACCA	2580
Qy	2581	TCCACCGTTGGTTTACACTGGTGTCCCTCCCAAGAGCCCAATGTTGAGATTGGAGTAT	2640
Db	2581	TCCACCGTTGGTTTACACTGGTGTCCCTCCCAAGAGCCCAATGTTGAGATTGGAGTAT	2640
Qy	2641	CTAATTCTTACAGCAACCAACACACCCAGCCCGGAAACACAGTGGACACTGCTCACCACA	2700
Db	2641	CTAATTCTTACAGCAACCAACACACCCAGCCCGGAAACACAGTGGACACTGCTCACCACA	2700

Qy	2701	GAGGGAACAATCTTTCAGTGCAGAGGTCCATGGCTTAGAGAGTGCACACTCGGTATTTCTTC	2760
Db	2701	GAGGGAACAATCTTTCAGTGCAGAGGTCCATGGCTTAGAGAGTGCACACTCGGTATTTCTTC	2760
Qy	2761	AAGATGGAGCCCGCAACAGAGTGGGCTTGGGCCCTTTTCCGCTTGCAGGATGTGATT	2820
Db	2761	AAGATGGAGCCCGCAACAGAGTGGGCTTGGGCCCTTTTCCGCTTGCAGGATGTGATT	2820
Qy	2821	ACTCTCAGAGACATTTCTCAGACTCTCTTGATGTGACGCGCTCAGGGCATCATCGTG	2880
Db	2821	ACTCTCAGAGACATTTCTCAGACTCTCTTGATGTGACGCGCTCAGGGCATCATCGTG	2880
Qy	2881	GGTGTCTGCTGGGCCCTTCTCTGCTCTCTGGCTGATGTGTGCTGCTTACGACAAAGC	2940
Db	2881	GGTGTCTGCTGGGCCCTTCTCTGCTCTCTGGCTGATGTGTGCTGCTTACGACAAAGC	2940
Qy	2941	TCCCAAGGAGCCCTTCTCCGATTTCTCTCAGGCAACCCAGGAAACCCAGCGCTC	3000
Db	2941	TCCCAAGGAGCCCTTCTCCGATTTCTCTCAGGCAACCCAGGAAACCCAGCGCTC	3000
Qy	3001	TACACAAGAGCTCGGCTTGGGCTTCCAGTGTCTCCAGTGTCTGCTGCTGAGTGGAGTCCCTC	3060
Db	3001	TACACAAGAGCTCGGCTTGGGCTTCCAGTGTCTCCAGTGTCTGCTGCTGAGTGGAGTCCCTC	3060
Qy	3061	GTGCATCTCTGCTCCCGAGGATTTGGTCCCAACACCTCAGATGTGGAAGCAAAAGGCTGAA	3120
Db	3061	GTGCATCTCTGCTCCCGAGGATTTGGTCCCAACACCTCAGATGTGGAAGCAAAAGGCTGAA	3120
Qy	3121	GTAACAAGCTTATGGTGGCAGTGTTCAGATTGCGGGGCCACTTCCAAAGAAAGATC	3180
Db	3121	GTAACAAGCTTATGGTGGCAGTGTTCAGATTGCGGGGCCACTTCCAAAGAAAGATC	3180
Qy	3181	TCCCTGGCTCAGCAGGGGGACCAACTGGGAGGCTCTGCTGCTCTCAGCGGCAACCGGG	3240
Db	3181	TCCCTGGCTCAGCAGGGGGACCAACTGGGAGGCTCTGCTGCTCTCAGCGGCAACCGGG	3240
Qy	3241	CAGGGTAGTGGTCCAAAGGCGGCTCTGAACCGCTGCTCTGCTGCTCTCAGCGGCAACCGGG	3300
Db	3241	CAGGGTAGTGGTCCAAAGGCGGCTCTGAACCGCTGCTCTGCTGCTCTCAGCGGCAACCGGG	3300
Qy	3301	CAGACACTGCTGTGCTCAAGCCCTGCTGTATGACGGCATTAAGAGCAACGGGAGAAAG	3360
Db	3301	CAGACACTGCTGTGCTCAAGCCCTGCTGTATGACGGCATTAAGAGCAACGGGAGAAAG	3360
Qy	3361	CGTCCCGCAGCTCAGGATCAGGTGGAGTGTGATGCTTCTCACTCCGACTTGGT	3420
Db	3361	CGTCCCGCAGCTCAGGATCAGGTGGAGTGTGATGCTTCTCACTCCGACTTGGT	3420
Qy	3421	GCATCCAAAGGATGCTCTGACCTCCACCTCCAAAGACTGGAGCCAGAGGAACCACTGACT	3480
Db	3421	GCATCCAAAGGATGCTCTGACCTCCACCTCCAAAGACTGGAGCCAGAGGAACCACTGACT	3480
Qy	3481	GCAGAGACTCTGCTTCCAGCTCTGAGCTGTGATCTGTCTCAAGSAGCAGACTGGCTG	3540
Db	3481	GCAGAGACTCTGCTTCCAGCTCTGAGCTGTGATCTGTCTCAAGSAGCAGACTGGCTG	3540
Qy	3541	GGCAGSAGCTGGGAGGGTGCACCAACCAACAGTGGGCGCAGAGGGCTCACTGGCTTG	3600
Db	3541	GGCAGSAGCTGGGAGGGTGCACCAACCAACAGTGGGCGCAGAGGGCTCACTGGCTTG	3600
Qy	3601	CCAGAGAGCAGTGGCTCTGCTCTGCTCTCAGACTCCAGCGCCAGCAGCTGTATAGAG	3660
Db	3601	CCAGAGAGCAGTGGCTCTGCTCTGCTCTCAGACTCCAGCGCCAGCAGCTGTATAGAG	3660
Qy	3661	GAGGCCCTGGGAAAGCTGCCAGCCCAAGCCCTGTGCTCTTAAAGTCAAGCTCAGCCCAAGC	3720
Db	3661	GAGGCCCTGGGAAAGCTGCCAGCCCAAGCCCTGTGCTCTTAAAGTCAAGCTCAGCCCAAGC	3720
Qy	3721	CTTCCAGGGCCCTGTCTCTCTGCTCAGGTCCCTTGCAGAGAGGAGATGGCTCA	3780
Db	3721	CTTCCAGGGCCCTGTCTCTCTGCTCAGGTCCCTTGCAGAGAGGAGATGGCTCA	3780

QY 3781 GGAAATGCAATGCGTACACATGTTGTGTACTAGAGATATCCATAGTCTTTGGAGC 3840
DB |||||
3781 GGAAATGCAATGCGTACACATGTTGTGTACTAGAGATATCCATAGTCTTTGGAGC 3840
QY |||||
3841 CTCCTTAGGGTCTTTGGCTGGGGTGGGAGAACTTTTACTCTCCCTCATATTTCTGCATCA 3900
DB |||||
3841 CTCCTTAGGGTCTTTGGCTGGGGTGGGAGAACTTTTACTCTCCCTCATATTTCTGCATCA 3900
QY |||||
3901 CATACAGGAGGACTTTGAGACACACAGCTCTGTGTAATGGACACAGCTGTGGAAGTCTGTGTGT 3960
DB |||||
3901 CATACAGGAGGACTTTGAGACACACAGCTCTGTGTAATGGACACAGCTGTGGAAGTCTGTGTGT 3960
QY |||||
3961 GT 4020
DB |||||
3961 GT 4020
QY |||||
4021 GCCTAGTTGACCTCTCGTGGCAGGATGGTGTAAACAGTGTATCAGTGCCAGCTCTTTTGAGCT 4080
DB |||||
4021 GCCTAGTTGACCTCTCGTGGCAGGATGGTGTAAACAGTGTATCAGTGCCAGCTCTTTTGAGCT 4080
QY |||||
4081 TTTAGCCTTGTACACTAGGCTTTTATTAACACTCTGAGAGTGTCTCCAGTGTGTGTCTAC 4140
DB |||||
4081 TTTAGCCTTGTACACTAGGCTTTTATTAACACTCTGAGAGTGTCTCCAGTGTGTGTCTAC 4140
QY |||||
4141 AAAGACAGCGCCAGCCCTCTTCTGT 4200
DB |||||
4141 AAAGACAGCGCCAGCCCTCTTCTGT 4200
QY |||||
4201 GGGCCTATGACACCGGAGCTACACAGATGGTGTGTATCCCTCGGCTCTTAAGTTC 4260
DB |||||
4201 GGGCCTATGACACCGGAGCTACACAGATGGTGTGTATCCCTCGGCTCTTAAGTTC 4260
QY |||||
4261 CAGATGTCTGGGTGAACCCAGCTCAGCTCCCTCTCTTTGAGCATCTCTGACCTAAAT 4320
DB |||||
4261 CAGATGTCTGGGTGAACCCAGCTCAGCTCCCTCTCTTTGAGCATCTCTGACCTAAAT 4320
QY |||||
4321 TTGTAATCTGGGAAGTGCCTGTGTGTGGGAAATCTTCTTTGCAACCTGTCTCTCTGCT 4380
DB |||||
4321 TTGTAATCTGGGAAGTGCCTGTGTGTGGGAAATCTTCTTTGCAACCTGTCTCTCTGCT 4380
QY |||||
4381 CTTCTCTTCAATTTGTTGTGGT 4440
DB |||||
4381 CTTCTCTTCAATTTGTTGTGGT 4440
QY |||||
4441 CTTCTCTTCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4500
DB |||||
4441 CTTCTCTTCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4500
QY |||||
4501 CTTGTCTTTATGTCTGT 4560
DB |||||
4501 CTTGTCTTTATGTCTGT 4560
QY |||||
4561 ACATGACTACCTCATGTCTGTCTGAGCATAGTGTGACCCCTGGGTCCCCACAGCTCCC 4620
DB |||||
4561 ACATGACTACCTCATGTCTGTCTGAGCATAGTGTGACCCCTGGGTCCCCACAGCTCCC 4620
QY |||||
4621 CTGCCAAACCGCTTCTCTGGGAGATGAGCCCACTCCAGTGTATCTGGGAAAGACCCCTTG 4680
DB |||||
4621 CTGCCAAACCGCTTCTCTGGGAGATGAGCCCACTCCAGTGTATCTGGGAAAGACCCCTTG 4680
QY |||||
4681 TGGCTTGTCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4740
DB |||||
4681 TGGCTTGTCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4740
QY |||||
4741 GAGTCAGGCTGCAAGGGGAAACCCCAAGAGGGGTAGGAGTGAACCAAGAGGCTGA 4800
DB |||||
4741 GAGTCAGGCTGCAAGGGGAAACCCCAAGAGGGGTAGGAGTGAACCAAGAGGCTGA 4800
QY |||||
4801 AAAAAATGGTGCACACCTCTGACACAGAGATGGGTGTGTCTTTTGAAGTGCAGTTC 4860
DB |||||
4801 AAAAAATGGTGCACACCTCTGACACAGAGATGGGTGTGTGTCTTTTGAAGTGCAGTTC 4860
QY |||||
4861 CTGGCTGAAATGAAAGGGGTGAGAGAGGGAGCTACTGGGGCTGCCATGGCTCAGTTCC 4920

DB 4861 CTGGCTGAAATGAAAGGGGTGAGAGAGGGAGCTACTGGGGCTGCCATGGCTCAGTTCC 4920
QY |||||
4921 CTGACCTTGAGACCTCGAACTGGCTTCAGAGTAGCAAGAGTTTCTTCCAAAGATGCTGT 4980
DB |||||
4921 CTGACCTTGAGACCTCGAACTGGCTTCAGAGTAGCAAGAGTTTCTTCCAAAGATGCTGT 4980
QY |||||
4981 AAGGGAAGTCTTTGTCATAGGAAAGGGCGGCTGGCTCATTTTATTTATCTTTCTTTTACA 5040
DB |||||
4981 AAGGGAAGTCTTTGTCATAGGAAAGGGCGGCTGGCTCATTTTATTTATCTTTCTTTTACA 5040
QY |||||
5041 CTGAATCCCAAAATCATCTTACCAAAAGGGCCAAAGCTGACTGGTATTTCTCTGAGTCAC 5100
DB |||||
5041 CTGAATCCCAAAATCATCTTACCAAAAGGGCCAAAGCTGACTGGTATTTCTCTGAGTCAC 5100
QY |||||
5101 AAGAGCCATGCACTCTCTGGTTTCTCACCTCAGTGTATGCTCCAGAAATGCTCAGATCCA 5160
DB |||||
5101 AAGAGCCATGCACTCTCTGGTTTCTCACCTCAGTGTATGCTCCAGAAATGCTCAGATCCA 5160
QY |||||
5161 GTGGCATCTGTGCTCTTGTGTCACATCTTTCTATTTCAACTGGCTGGGCACATCAAGTGT 5220
DB |||||
5161 GTGGCATCTGTGCTCTTGTGTCACATCTTTCTATTTCAACTGGCTGGGCACATCAAGTGT 5220
QY |||||
5221 AACTCTGGCTTCTGGGCCAAGTTAGAAATAACAGTCTATTTTCCCTTTTATTTTATTTA 5280
DB |||||
5221 AACTCTGGCTTCTGGGCCAAGTTAGAAATAACAGTCTATTTTCCCTTTTATTTTATTTA 5280
QY |||||
5281 TTTTATTTTATTTTATGCTTTTCTAGTGGAGTGTAGTCTCTGAAAGCGTCTGTGTTTAT 5340
DB |||||
5281 TTTTATTTTATTTTATGCTTTTCTAGTGGAGTGTAGTCTCTGAAAGCGTCTGTGTTTAT 5340
QY |||||
5341 AGCCTTGTGTCTCACTCATGTTTGAACCCCAACCATTTCTCTCTCTCTCTCTCTCTCTCAGC 5400
DB |||||
5341 AGCCTTGTGTCTCACTCATGTTTGAACCCCAACCATTTCTCTCTCTCTCTCTCTCTCAGC 5400
QY |||||
5401 CAGCCTATGATAACACTAAAGATTTAAATGCTGGCTTGTGATCTCATTAAGACAGAT 5460
DB |||||
5401 CAGCCTATGATAACACTAAAGATTTAAATGCTGGCTTGTGATCTCATTAAGACAGAT 5460
QY |||||
5461 TGTCACTTGAATCTTCTATAGCATTTCAAGTGGCCACGGCCACACACCGATGTTT 5520
DB |||||
5461 TGTCACTTGAATCTTCTATAGCATTTCAAGTGGCCACGGCCACACACCGATGTTT 5520
QY |||||
5521 CTTCAATGCTCTGAAAGTCAAGAGCTCATTTTGTGTTTCTCTGGTTAGATTTCTTTCTCTCC 5580
DB |||||
5521 CTTCAATGCTCTGAAAGTCAAGAGCTCATTTTGTGTTTCTCTGGTTAGATTTCTTTCTCTCC 5580
QY |||||
5581 TTGCTTCAATGAAATAACCGTTTAAACAGTAGGCTTTAGCATCACACCATAGTCAAT 5640
DB |||||
5581 TTGCTTCAATGAAATAACCGTTTAAACAGTAGGCTTTAGCATCACACCATAGTCAAT 5640
QY |||||
5641 TCCTCATGTTCTTGTGTTTAAACAGCATTTGAGGTTCTGGGTTTAAATTAATAGCTGCAAA 5700
DB |||||
5641 TCCTCATGTTCTTGTGTTTAAACAGCATTTGAGGTTCTGGGTTTAAATTAATAGCTGCAAA 5700
QY |||||
5701 TGAGACAAATTTATTAACCAATTTAGGCTGGGTGGAATAATGTTCTCAAAAGCAATTAAGTAA 5760
DB |||||
5701 TGAGACAAATTTATTAACCAATTTAGGCTGGGTGGAATAATGTTCTCAAAAGCAATTAAGTAA 5760
QY |||||
5761 TAAATCTGGTATCTGCTTATTAACCTCAAGTGTGATTAAGAAAGTAGCAGAACTCACTAGCA 5820
DB |||||
5761 TAAATCTGGTATCTGCTTATTAACCTCAAGTGTGATTAAGAAAGTAGCAGAACTCACTAGCA 5820
QY |||||
5821 TTATATATGATTTGGGTTCTGAGTAACTGGGGAGTGTAGCTTTGTGTGACTTTGTAGCACC 5880
DB |||||
5821 TTATATATGATTTGGGTTCTGAGTAACTGGGGAGTGTAGCTTTGTGTGACTTTGTAGCACC 5880
QY |||||
5881 AGGTCTCTTATAGGAAAGTCTGTGTGGCTTTTACAGGCAATAGTCCCTTTGTCTGTTGCC 5940
DB |||||
5881 AGGTCTCTTATAGGAAAGTCTGTGTGGCTTTTACAGGCAATAGTCCCTTTGTCTGTTGCC 5940
QY |||||
5941 ATGATGCTCTTAAAGTCTCTTTGGAGTCTCATTTTAAAGATTTCTTTTCTCGAAGCATGACAA 6000

Db 5941 ATGATGCTTAAGTCTCTCGAGTCTCATTTAAGAAATTCCTTTCTCGAAGCATGACAA 6000
Qy 6001 GTGTATCGCAATATCTTACATGCTCACTGCTTTTACCTGCTAGTTTGTGCTGGGTATTT 6060
Db 6001 GTGTATCGCAATATCTTACATGCTCACTGCTTTTACCTGCTAGTTTGTGCTGGGTATTT 6060
Qy 6061 AATTGCACTTTCCAGCATGCTGCTGCTTCCCTTACAAATATGATATCTTTATTTTACAC 6120
Db 6061 AATTGCACTTTCCAGCATGCTGCTGCTTCCCTTACAAATATGATATCTTTATTTTACAC 6120
Qy 6121 TAAGTGTGTGATGATGATGCTGCTGCTTAAAGATTAATAAATTTTCCAGAC 6176
Db 6121 TAAGTGTGTGATGATGATGCTGCTGCTTAAAGATTAATAAATTTTCCAGAC 6176

RESULT 2
AF176694 6176 bp mRNA linear ROD 27-APR-2000
LOCUS Mus musculus neighbor of Punc ell protein (Nope) mRNA, complete
cde.
ACCESSION AF176694
VERSION AF176694.1 GI:7650185
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS
TITLE Cloning and expression of nope, a new mouse gene of the immunoglobulin superfamily related to guidance receptors
JOURNAL Genomics 64 (1), 15-23 (2000)
MEDLINE 20175427
PUBMED 10708514
REFERENCE 2 (bases 1 to 6176)
AUTHORS Salbaum, J.M.
TITLE Direct Submission
JOURNAL Submitted (09-AUG-1999) The Neurosciences Institute, 10640 John Jay Hopkins Drive, San Diego, CA 92121, USA

FEATURES
source
1. .6176
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB"
/db_xref="taxon:10090"
/chromosome="9"
/map="band D-E1; 24.59 cR from D9Mit143"
1. .6176
/gene="Nope"
1. .3759
/gene="Nope"
/note="immunoglobulin superfamily; related to guidance receptors"
/codon_start=1
/product="neighbor of Punc ell protein"
/protein_id="AAF65930.1"
/db_xref="GI:7650186"
/translation="MARADTGRGLLVITPCLLSARGELPLIPEQTTVKLSDEBPLQVI LGPEQAVVWCTLGATAGPRTVTWSKGDVLEHENLHLLPNGLSPLSEBDS DDEALRIWKTGYSGLSPHSLGVASQVAVVVKLATLEDLSLHPSQIVBENGSTAR FECHTGLPAPIITWEKQVTVPEPRLLITLPMKLLQILDVQSDAGSYRCVATNSAR QRFQASLTVALRGLSEATRGQDVIVAAPENTTVVSGQNVMECVASADPTPFVSM VRQKGPISDVIIVLGRNLLIASAQRHSGVYCRANKPLTRDPATAAAEALRLVAP AISQPSALSKRSTASTARFVCRASGEPFPAHLWHDGIPLPNRRGRVKKVGGGSLVIT QIGQDAGYQCAENASAGTACAAAPLAVVVRGLSPAPRTVATPILSSSLVWAMER PELHSEQIIQFSLHYQKARGVNVYQFVAVNDDTELQVRDLEPTDPTFYVAYSOL GASTSSPALTLDVPSAPOLITLSPNPSDIRVAWLPLPSLSNGVLYKXIYEG LGKEDVPSTVEPCNETQLTLSLQPNKYVRVIRISACTGAGYGVPSOMHRTFVGN QSHVPAPAEVLKRAKESLVSVSPHPPTQISGKLYWGEVTEAEADGRPPGR GDQAWDPVRLKKKXQYELTQVPGRPYELVAFNKHEDYAAVWKGTEKATP DLPTQRPGLPAPAHVAESNSTSIWLRWKKPFTTVKIVNYTVRFGPWGLRNASLVT YTSSTSGDIIIGLKLKPTKYEFAVQSHGVMDGPFSGVVERSTLPDRPSPDLRLS PLTSTVRLHWCPTPEPGEIVEYLILYSNNHTQPEHQWTLILLTEGNIIFSAEVHGLSS

misc_feature
1. .23
/gene="Nope"
/note="unable to sequence region in mRNA"

BASE COUNT 1316 a 1756 c 1634 g 1470 t
ORIGIN
Query Match 100.0%; Score 6176; DB 10; Length 6176;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6176; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGCGGGGGGACACGGGGCGGGGCTCTGGTGTGTGACCTTCTGCTGCTGCTGCGG 60
Db 1 ATGGCGGGGGGACACGGGGCGGGGCTCTGGTGTGTGACCTTCTGCTGCTGCTGCGG 60
Qy 61 CGGGGGGAGCTGCCATTGCCCCCAGGAGACAACCTGTCAAGCTGAGCTGTGATGAGGACCC 120
Db 61 CGGGGGGAGCTGCCATTGCCCCCAGGAGACAACCTGTCAAGCTGAGCTGTGATGAGGACCC 120
Qy 121 CTGCAAGTGTATCTGGGCCCTGAGCAGGCTGTGTGTGCTGGACTGCACTTTGGGGGCTACA 180
Db 121 CTGCAAGTGTATCTGGGCCCTGAGCAGGCTGTGTGTGCTGGACTGCACTTTGGGGGCTACA 180
Qy 181 GCTGCTGGGCTCGGACAGGCTGACATGAGCAGGATGAGACACTGTACTAGAGCAT 240
Db 181 GCTGCTGGGCTCGGACAGGCTGACATGAGCAGGATGAGACACTGTACTAGAGCAT 240
Qy 241 GAGAACCTGCACCTGTACCCCAATGGCTCTCTGTGGCTGCTCACCCCTAGACAGAA 300
Db 241 GAGAACCTGCACCTGTACCCCAATGGCTCTCTGTGGCTGCTCACCCCTAGACAGAA 300
Qy 301 GACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTCTGT 360
Db 301 GACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTCTGT 360
Qy 361 CTGCCCCACAGCCCGCTAGGAGTGTGGCCAGCCAGGTGTGCTGTTGCTCAAGCTTGCCACA 420
Db 361 CTGCCCCACAGCCCGCTAGGAGTGTGGCCAGCCAGGTGTGCTGTTGCTCAAGCTTGCCACA 420
Qy 421 CTGGAACACTTCTCTGCAACCCGAGTCCAGAGTCTGAGGAGGAGACGGACAGCAGC 480
Db 421 CTGGAACACTTCTCTGCAACCCGAGTCCAGAGTCTGAGGAGGAGACGGACAGCAGC 480
Qy 481 TTTGAATGCCACCAAGGGCTTTCCAGCCCCCATCAATTAATTTGGGAAAAGGACAGGTG 540
Db 481 TTTGAATGCCACCAAGGGCTTTCCAGCCCCCATCAATTAATTTGGGAAAAGGACAGGTG 540
Qy 541 ACCGTGCTGAGGAGCCCCCGCTCATCATCTTTCCCAAGTGGCTCTCTCCAGATCCTAGAT 600
Db 541 ACCGTGCTGAGGAGCCCCCGCTCATCATCTTTCCCAAGTGGCTCTCTCCAGATCCTAGAT 600
Qy 601 GTCCAGACAGTGTGATGAGGCTCTCAAGCTGTGGTGGCCCAATTCAGCCCCGCAACGA 660
Db 601 GTCCAGACAGTGTGATGAGGCTCTCAAGCTGTGGTGGCCCAATTCAGCCCCGCAACGA 660
Qy 661 TTCAAGCAGAGGGCTCGCTCACTGTGGCCCTCAGAGGGCTCTTTGGAGGGCTACCAAGGGG 720
Db 661 TTCAAGCAGAGGGCTCGCTCACTGTGGCCCTCAGAGGGCTCTTTGGAGGGCTACCAAGGGG 720
Qy 721 CAGGATGTGTCAATTTGGGAGCCCCCAGAGAACCAACCGGTAGTGTCTGGACAGAAATGTA 780
Db 721 CAGGATGTGTCAATTTGGGAGCCCCCAGAGAACCAACCGGTAGTGTCTGGACAGAAATGTA 780
Qy 781 GTGATGAGTGTGCTGGGCTCTGCTGACCCCAACCCCTTTTGTGTCTCTGGTCCGACAGAT 840
Db 781 GTGATGAGTGTGCTGGGCTCTGCTGACCCCAACCCCTTTTGTGTCTCTGGTCCGACAGAT 840
Qy 841 GGAAAGCTTATCTCCAGGATGTTCATGTTCTGGGGCCGACCAATCTACTCATCGCCAGC 900

Qy 3061 GTGCATCTCTGTCCTCCAGAGATTGTTGCCCAACACCTCTAGATGTGGAGAACAGGCTGAA 3120
Db GTGCATCTCTGTCCTCCAGAGATTGTTGCCCAACACCTCTAGATGTGGAGAACAGGCTGAA 3120
Qy 3121 GTACACAGCCTTATGGGTGGCAGTGTTCAGATTGCCGGGCCACTCCAGAGAAAGTC 3180
Db GTACACAGCCTTATGGGTGGCAGTGTTCAGATTGCCGGGCCACTCCAGAGAAAGTC 3180
Qy 3181 TCCTGGGCTCAGGCAAGGGGAGACCAAACTGGGCAGGCTCTGGGCAGGCTGTGAGCTGCC 3240
Db TCCTGGGCTCAGGCAAGGGGAGACCAAACTGGGCAGGCTCTGGGCAGGCTGTGAGCTGCC 3240
Qy 3241 CAGGCTAGTGTCCAAAGCCCGGCTCTGACCCGCTCTGACCCGCTCTGCTGCTCCAGCGGAAACCGG 3300
Db CAGGCTAGTGTCCAAAGCCCGGCTCTGACCCGCTCTGACCCGCTCTGCTGCTCCAGCGGAAACCGG 3300
Qy 3301 CAGACATGCTGCTGCAAGCCCTGGTGTATGACGGCATAAAGAGCAACGGGAGAAAG 3360
Db CAGACATGCTGCTGCAAGCCCTGGTGTATGACGGCATAAAGAGCAACGGGAGAAAG 3360
Qy 3361 CCGTCCCAGGCTGCAGGAATCAGGTGGAAGCTCAGGTCTATGTCCTCCACTCCGACTTCGGT 3420
Db CCGTCCCAGGCTGCAGGAATCAGGTGGAAGCTCAGGTCTATGTCCTCCACTCCGACTTCGGT 3420
Qy 3421 GCATCCAAAGATGTCCTGACCTCCACTCCAAAGCTGGAGCCAGAGGAACCACTGACT 3480
Db GCATCCAAAGATGTCCTGACCTCCACTCCAAAGCTGGAGCCAGAGGAACCACTGACT 3480
Qy 3481 GCAGAGCTCTGCTTCCAGCTCTGGAGCTGTGATCTGCTCAAGAGCAGACTGGCTG 3540
Db GCAGAGCTCTGCTTCCAGCTCTGGAGCTGTGATCTGCTCAAGAGCAGACTGGCTG 3540
Qy 3541 GGCAGGGAGCTGGAGGGTGCACCAACCAACAGTGGGCCAGAGAGGCTCACCTGCTTG 3600
Db GGCAGGGAGCTGGAGGGTGCACCAACCAACAGTGGGCCAGAGAGGCTCACCTGCTTG 3600
Qy 3601 CCAGAAGCAGCAGTGGCTCTGCTGCTGCTCAGACCTCCAGCCAGCAGCTGTATAGAG 3660
Db CCAGAAGCAGCAGTGGCTCTGCTGCTGCTCAGACCTCCAGCCAGCAGCTGTATAGAG 3660
Qy 3661 GAGGCCCTGGGAAAGCTGCGCCAGCCCAAGCCCTGTGCTCTTAAAGTCAAGCCCAAGC 3720
Db GAGGCCCTGGGAAAGCTGCGCCAGCCCAAGCCCTGTGCTCTTAAAGTCAAGCCCAAGC 3720
Qy 3721 CTTCCAGGGCCCTGCTCTCTGCTCAGTGTCCCTGAGCAGAGGAGGAGATATGGCTCA 3780
Db CTTCCAGGGCCCTGCTCTCTGCTCAGTGTCCCTGAGCAGAGGAGGAGATATGGCTCA 3780
Qy 3781 GGAACATGCCATGCATGCTACACATGTGTGTACTAGAGATATCCATAAGTCTCTTGGAGC 3840
Db GGAACATGCCATGCATGCTACACATGTGTGTACTAGAGATATCCATAAGTCTCTTGGAGC 3840
Qy 3841 CTTTAGGGTCTTTGGCTGGGTTGGGAGAACTTTTACTCTCCCTCATATTTCTGCAATCA 3900
Db CTTTAGGGTCTTTGGCTGGGTTGGGAGAACTTTTACTCTCCCTCATATTTCTGCAATCA 3900
Qy 3901 CATACAGAGGACCTTGAGACACAGCTCTGTGTATGAGACACGCTGTGAAGTCTGTGTGT 3960
Db CATACAGAGGACCTTGAGACACAGCTCTGTGTATGAGACACGCTGTGAAGTCTGTGTGT 3960
Qy 3961 GT 4020
Db GT 4020
Qy 4021 GCCTAGTTGACCCCTCCGTCAGGATGTGTAAACAGTGTATCAGTGCAGCTCTTTGAGCT 4080
Db GCCTAGTTGACCCCTCCGTCAGGATGTGTAAACAGTGTATCAGTGCAGCTCTTTGAGCT 4080
Qy 4081 TTTAGCCTGTGCTAGCTCTTTTATACATCTGAGAGTGTCTCCAGTGTGTGTCTAC 4140
Db TTTAGCCTGTGCTAGCTCTTTTATACATCTGAGAGTGTCTCCAGTGTGTGTCTAC 4140

Qy 4141 AAAGACAGCGCCCGAGCCCTCTTCTGTGAGCTGTGCTGAGCAGAGTGCAGTCACTCCAC 4200
Db AAAGACAGCGCCCGAGCCCTCTTCTGTGAGCTGTGCTGAGCAGAGTGCAGTCACTCCAC 4200
Qy 4201 GGGCTATGACACCGCAGCCTTACACAGCATGCTGTATCCCTGGCTCTTCAAGGTC 4260
Db GGGCTATGACACCGCAGCCTTACACAGCATGCTGTATCCCTGGCTCTTCAAGGTC 4260
Qy 4261 CAGATGTCTGGGTGAACCCAGCTCAGCTCCCTCTCTCTTTGAGCATCTCTGTACCTAAT 4320
Db CAGATGTCTGGGTGAACCCAGCTCAGCTCCCTCTCTCTTTGAGCATCTCTGTACCTAAT 4320
Qy 4321 TTGTAATCTGGGAAGTGCCTGGTTGGGAAATCTTCTTTTGCACCCCTGTCTCTGTC 4380
Db TTGTAATCTGGGAAGTGCCTGGTTGGGAAATCTTCTTTTGCACCCCTGTCTCTGTC 4380
Qy 4381 CTTTCTCTCATTTTCTCTGCTGTCTCATGTCTATCTTGTCTGATATCTCTGGGCT 4440
Db CTTTCTCTCATTTTCTCTGCTGTCTCATGTCTATCTTGTCTGATATCTCTGGGCT 4440
Qy 4441 CTTTCTCTCTCCATGATGCCCCCTGATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4500
Db CTTTCTCTCTCCATGATGCCCCCTGATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4500
Qy 4501 CTTGTCTTTATGCTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4560
Db CTTGTCTTTATGCTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4560
Qy 4561 ACATGACTACTCATGTCT 4620
Db ACATGACTACTCATGTCT 4620
Qy 4621 CTGCAACCCGCTTCTCTGGCAGATGAGCCCACTCTCAAGTAGATCTCTGAAAAGACCTTG 4680
Db CTGCAACCCGCTTCTCTGGCAGATGAGCCCACTCTCAAGTAGATCTCTGAAAAGACCTTG 4680
Qy 4681 TGGCTTGTCTGGCTGCCCTCCCTTGGTGTGATGAGAGTGTCTATGGAAGAGAT 4740
Db TGGCTTGTCTGGCTGCCCTCCCTTGGTGTGATGAGAGTGTCTATGGAAGAGAT 4740
Qy 4741 GAGTCCAGGCTGCACAGGGGAAACCCCAAGAGGGTAGGGTAGAACCAAGAGGCTGA 4800
Db GAGTCCAGGCTGCACAGGGGAAACCCCAAGAGGGTAGGGTAGAACCAAGAGGCTGA 4800
Qy 4801 AAAAAATGCTGCTCCACCCCTCTGCAAGAGAGATGGGTGTGCTTTTGAAGTCTGAGTC 4860
Db AAAAAATGCTGCTCCACCCCTCTGCAAGAGAGATGGGTGTGCTTTTGAAGTCTGAGTC 4860
Qy 4861 CTGCTCAAACTGAAGGGGTGAGAGAGGGAGCTACTGGGGCTGCAATGGCTCAGTTTC 4920
Db CTGCTCAAACTGAAGGGGTGAGAGAGGGAGCTACTGGGGCTGCAATGGCTCAGTTTC 4920
Qy 4921 CTGACCTGGAGCCCTGAACTGGCTTCAAGTAGCAAGAGTGTCTCTCAAGATGCTGT 4980
Db CTGACCTGGAGCCCTGAACTGGCTTCAAGTAGCAAGAGTGTCTCTCAAGATGCTGT 4980
Qy 4981 AAGGGAAGTCTTTGTCATAGGAAAGGGCGGCTGGCTCATTTTATTTTCTTTTACA 5040
Db AAGGGAAGTCTTTGTCATAGGAAAGGGCGGCTGGCTCATTTTATTTTCTTTTACA 5040
Qy 5041 CTGAATCCCAAAATCATCTTACCAAGGGCCAAAGCTGATGCTGGTATTTCTCAGTCCAC 5100
Db CTGAATCCCAAAATCATCTTACCAAGGGCCAAAGCTGATGCTGGTATTTCTCAGTCCAC 5100
Qy 5101 AAGAGCCTAGCCATCTCTCTGGTTTCTCACTCAGTCTATGCTCCAGAAATGTCAGATCCA 5160
Db AAGAGCCTAGCCATCTCTCTGGTTTCTCACTCAGTCTATGCTCCAGAAATGTCAGATCCA 5160
Qy 5161 GTGCACTGTGCTCTTGTGCACTCTTTCTATTTCAAATGGGCTGCAATCAAGTGT 5220
Db GTGCACTGTGCTCTTGTGCACTCTTTCTATTTCACTTGGCTGCAATCAAGTGT 5220
Qy 5221 AACTCTGGCTTCTGGGCCAAAGTTAGAAATAACCAAGTCTATTTTCTCTTTTATTTT 5280

Qy 121 CTGCAAGTGTATCTTGGGCGCTTGAGCAGGCTGTGTGTGCTGAGTGTGACTGTGACATTTTGGGGCTTACA 180
Db 170 CTGCAAGTGTATCTTGGGCGCTTGAGCAGGCTGTGTGTGCTGAGTGTGACTGTGACATTTTGGGGCTTACA 229
Qy 181 GCTCTGGGCTTCCGACCAAGGCTGACATGGAGCAAGTGTGAGACACTGTACTAGAGCAT 240
Db 230 GCTCTGGGCTTCCGACCAAGGCTGACATGGAGCAAGTGTGAGACACTGTACTAGAGCAT 289
Qy 241 GAGAACTGTGACCTGTACCAATGGCTCCCTGTGGCTGTCTTCCACCCCTAGAGCAAGAA 300
Db 290 GAGAACTGTGACCTGTACCAATGGCTCCCTGTGGCTGTCTTCCACCCCTAGAGCAAGAA 349
Qy 301 GACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTTCCTGT 360
Db 350 GACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTTCCTGT 409
Qy 361 CTGGCCACAGCCGCTTAGGAGTGGTGGCCAGCCAGGTGTGTGTGCTCAAGCTTGGCCACA 420
Db 410 CTGGCCACAGCCGCTTAGGAGTGGTGGCCAGCCAGGTGTGTGTGCTCAAGCTTGGCCACA 469
Qy 421 CTCGAAGACTTCTCTCTGACCCCGAGTCCAGATTGTGGAGGAGAACGGGACAGCAGC 480
Db 470 CTCGAAGACTTCTCTCTGACCCCGAGTCCAGATTGTGGAGGAGAACGGGACAGCAGC 529
Qy 481 TTTGAATGCCACACCAAGGGCTTCCAGCCCCCATCATTTACTTTGGGAAAAGGACCAAGTG 540
Db 530 TTTGAATGCCACACCAAGGGCTTCCAGCCCCCATCATTTACTTTGGGAAAAGGACCAAGTG 589
Qy 541 ACCGTGCTGAGGAGCCCGGCTCATCATCTTCCCAAGTGGCTCTTCCAGATCCTAGAT 600
Db 590 ACCGTGCTGAGGAGCCCGGCTCATCATCTTCCCAAGTGGCTCTTCCAGATCCTAGAT 649
Qy 601 GTCAGAGACATGATGAGGCTCTTACCGCTGGTGGCCACCAATTCAGCCCGCCCAAGCA 660
Db 650 GTCAGAGACATGATGAGGCTCTTACCGCTGGTGGCCACCAATTCAGCCCGCCCAAGCA 709
Qy 661 TTTAGCCAGGAGGCTCGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTTACAGAGGG 720
Db 710 TTTAGCCAGGAGGCTCGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTTACAGAGGG 769
Qy 721 CAGGATGTGTCTATTTGGGAGCCCGCAGAGAACCAACGGTGTGTCTGGACAGAAATGTA 780
Db 770 CAGGATGTGTCTATTTGGGAGCCCGCAGAGAACCAACGGTGTGTCTGGACAGAGTGT 829
Qy 781 GTGATGGAGTGGTGGCTCTGTGACCCCGCTTTTGTGTCTTGGGCTCCGACAGAT 840
Db 830 GTGATGGAGTGGTGGCTCTGTGACCCCGCTTTTGTGTCTTGGGCTCCGACAGAT 889
Qy 841 GGAAGCCTATCTCCAGGATGTCTATCGTTCTGGGCGGACCAATCTACTCATCGCCAGC 900
Db 890 GGAAGCCTATCTCCAGGATGTCTATCGTTCTGGGCGGACCAATCTACTCATCGCCAGC 949
Qy 901 GCGCAGCCTCGGCACTCTGAGTCTATGTCTGCGGAGCCAAACAGCCCTTCAAGCGTGAC 960
Db 950 GCGCAGCCTCGGCACTCTGAGTCTATGTCTGCGGAGCCAAACAGCCCTTCAAGCGTGAT 1009
Qy 961 TTGCGCACTGGGCTGTGAGCTCGAGTGTGTGTCGCCAGCCATCTCGCAGGACCC 1020
Db 1010 TTGCGCACTGGGCTGTGAGCTCGAGTGTGTGTCGCCAGCCATCTCGCAGGACCC 1069
Qy 1021 GAGGCGCTCTCGCGAGCGGGCCAGCACCGCGCTTCTGTGTGCGGGCTTCCGGGGAG 1080
Db 1070 GAGGCGCTCTCGCGAGCGGGCCAGCACCGCGCTTCTGTGTGCGGGCTTCCGGGGAG 1129
Qy 1081 CCAGGGCCCGGCTGCACTGGCTGCAAGACGGGATCCCGTTGGACCAATAGGGCGGCTC 1140
Db 1130 CCAGGGCCCGGCTGCACTGGCTGCAAGACGGGATCCCGTTGGACCAATAGGGCGGCTC 1189
Qy 1141 AAGGTGACGGCGGTGGCGGAGCTTGGTTCATCACTCAGATCGGCTGACGAGACGCTGGC 1200
Db 1190 AAGGTGACGGCGGTGGCGGAGCTTGGTTCATCACTCAGATCGGCTTTCAGGACGCTGGC 1249

Qy 1201 TACTACAGTGTGTAGCAGAAACAGCGCGGGAACTGCTGTGTGCTGCGCTGCGCCCTTGGCG 1260
Db 1250 TACTACAGTGTGTGTAGCAGAAACAGCGCGGGAACTGCTGTGTGCTGCGCTGCGCCCTTGGCG 1309
Qy 1261 GTAGTGTGTGCGCAGGGGCTGCGCAGCGCCCGACTCGGGTCAAGCCACGCGCTGAGC 1320
Db 1310 GTAGTGTGTGCGCAGGGGCTGCGCAGCGCCCGACTCGGGTCAAGCCACGCGCTGAGC 1369
Qy 1321 AGCTCTCTCTGTGTGTGCGCTGCGGAGCGGCTGAGTTGACAGCGAGCAAAATCATTTGGC 1380
Db 1370 AGCTCTCTCTGTGTGTGCGCTGCGGAGCGGCTGAGTTGACAGCGAGCAAAATCATTTGGC 1429
Qy 1381 TTCTCTCTTCACTACCAAAAGGCAAGGGAGTGGACAATGTGAGTACCAAGTTTGCAGTA 1440
Db 1430 TTCTCTCTTCACTACCAAAAGGCAAGGGAGTGGACAATGTGAGTACCAAGTTTGCAGTA 1489
Qy 1441 AACAAATGACACACAGAGCTGCGGAGCTTGGGAGCTTGGAACTTGGAACTTGGATTC 1500
Db 1490 AACAAATGACACACAGAGCTGCGGAGCTTGGGAGCTTGGAACTTGGAACTTGGATTC 1549
Qy 1501 TAGCTGTGTGCTTACTTCCAGCTGGGGCCAGCGCAACTTCCAGCCCGAGCCCTTGGTGCAT 1560
Db 1550 TAGCTGTGTGCTTACTTCCAGCTGGGGCCAGCGCAACTTCCAGCCCGAGCCCTTGGTGCAT 1609
Qy 1561 ACATGGAACGATGTCTCCAGCGCAGCACCCCAAGCTTACTTGTCTCAGCCCCCAACCCCTCG 1620
Db 1610 ACATGGAACGATGTCTCCAGCGCAGCACCCCAAGCTTACTTGTCTCAGCCCCCAACCCCTCG 1669
Qy 1621 GACATCAGGCTGCGATGGCTGCGCTGCTCAGCTCAGCTGAGCAATGAGCAGGTGTGAAG 1680
Db 1670 GACATCAGGCTGCGATGGCTGCGCTGCTCAGCTCAGCTGAGCAATGAGCAGGTGTGAAG 1729
Qy 1681 TACAAGATAGAGTACGGTTTGGGGAGGAGATCAGGTTTTTCTCCACCGAGGTGCGCTGGA 1740
Db 1730 TACAAGATAGAGTACGGTTTGGGGAGGAGATCAGGTTTTTCTCCACCGAGGTGCGCTGGA 1789
Qy 1741 AATGAGACACACTTACGTTAAACTCACTTTCAGCCAAACAAAGTGTACCGAGTCCGAT 1800
Db 1790 AATGAGACACACTTACGTTAAACTCACTTTCAGCCAAACAAAGTGTACCGAGTCCGAT 1849
Qy 1801 TCAGCTGGCACTGCGCTGTGTATGAGTCTCTTCTCAGTGTGATGAGCAGAGACACT 1860
Db 1850 TCAGCTGGCACTGCGCTGTGTATGAGTCTCTTCTCAGTGTGATGAGCAGAGACACT 1909
Qy 1861 GGTGTGCACAAACAGAGCCATGTTCCCTTTGCCCCCTGCAAAATGGAAGGTGAGGGCAAG 1920
Db 1910 GGTGTGCACAAACAGAGCCATGTTCCCTTTGCCCCCTGCAAAATGGAAGGTGAGGGCAAG 1969
Qy 1921 ATGAGGTCCCTGTGTGTGTATGCGAGCGCCCGCTTCCAGCCAGAGATCTCTTGGATAC 1980
Db 1970 ATGAGGTCCCTGTGTGTGTATGCGAGCGCCCGCTTCCAGCCAGAGATCTCTTGGATAC 2029
Qy 1981 AAACTCTTACTGGGGAGAGTGGGAAACAGAGAGGAGGAGCAGATGATGACCGCCCCCAGGG 2040
Db 2030 AAACTCTTACTGGAGAGAGTGGGAAACAGAGAGGAGGAGCAGATGATGACCGCCCCCAGGG 2089
Qy 2041 GGTGTGGAGATCAAGCTTGGGAGCTGCGGCGCGTGTGAGAGAGAGTGAAGCAG 2100
Db 2090 GGTGTGGAGATCAAGCTTGGGAGCTGCGGCGCGTGTGAGAGAGAGTGAAGCAG 2149
Qy 2101 TATGAACCTGACCCACTTGTAGTCCCTGGCAGCGCCGTATCAGAGTGAAGCTCTGTAGCTTCAAC 2160
Db 2150 TATGAACCTGACCCACTTGTAGTCCCTGGCAGCGCTGTATCAGAGTGAAGCTCTGTAGCTTCAAC 2209
Qy 2161 AAAACAAGAGCAGCTACGCTGTGTGTGAAAGGGCAAGACGGAGAGGGCGCCACGCCA 2220
Db 2210 AAAACAAGAGCAGCTACGCTGTGTGTGAAAGGGCAAGACGGAGAGGGCGCCACGCCA 2269
Qy 2221 GACCTGCTATCCAGAGGGGCGCACCGCTGCTCTGCGCCATGTCTCAGCAGAGTCAAC 2280
Db 2270 GACCTGCTATCCAGAGGGGCGCACCGCTGCTCTGCGCCATGTCTCAGCAGAGTCAAC 2329
Qy 2281 AGCTTCACTTCCATTTGGCTTGTGTGAGAGACGACACTTTACCACTGTCAAGATTGTC 2340

Qy 4501 CTGTGCTTTATGTGTGTGTTCTGTCCTCCAGAGTTCAACCTATGACCCCTTTCTTAACA 4560
Db 4550 CTGTGCTTTATGTGTGTGTTCTGTCCTCCAGAGTTCAACCTATGACCCCTTTCTTAACA 4609
Qy 4561 ACATGACTACCTCATGTCTGCTTCAGACCATAGTGTGACCCCTGGGTCCCAAGAGCTCCC 4620
Db 4610 ACATGACTACCTCATGTCTGCTTCAGACCATAGTGTGACCCCTGGGTCCCAAGAGCTCCC 4669
Qy 4621 CTGCCAACCCTCTCTCGGCGAGATGAGCCCACTCCAAGTAGATCTGGAAGAGACCCCTTG 4680
Db 4670 CTGCCAACCCTCTCTCGGCGAGATGAGCCCACTCCAAGTAGATCTGGAAGAGACCCCTTG 4727
Qy 4681 TGGCTTGTCTGCTGCTCCCTCCCTGTTGTTGAGATGAGAGGTTTCTATGGAAGAGAT 4740
Db 4728 TGGCTTGTCTGCTGCTCCCTCCCTGTTGTTGAGATGAGAGGTTTCTATGGAAGAGAT 4787
Qy 4741 GAGTCCAGGCTGCAAGGGGACCCCAAGAGGGGTAGGAGTGAACCAAGAGGCTGA 4800
Db 4788 GAGTCCAGGCTGCAAGGGGACCCCAAGAGGGGTAGGAGTGAACCAAGAGGCTG- 4846
Qy 4801 AAAAAAATGGCTGCCACCCCATCTGCACAGAGAGATGGGTGTGCTTTTGACGTGCAGTC 4860
Db 4847 AAAAAAATGGCTGCCACCCCATCTGCACAGAGAGATGGGTGTGCTTTTGACGTGCAGTC 4906
Qy 4861 CTGGCTGAAAATGAAAGGGGTGAGGAGAGGGAGCTACTGGGGCTGCCATGGCTCAGTTTC 4920
Db 4907 CTGGCTGAAAATGAAAGGGGTGAGGAGAGGGAGCTACTGGGGCTGCCATGGCTCAGTTTC 4966
Qy 4921 CTGACCTGGAGCCCTGAACCTGGCTTCAGAGTAGCAAGAGTTTCCTCCAGAGTGTGT 4980
Db 4967 CTGACCTGGAGCCCTGAACCTGGCTTCAGAGTAGCAAGAGTTTCCTCCAGAGTGTGT 5026
Qy 4981 AAGGGAAGTCTTTGCATAGGAAAAGGGCGGCTGGCTCATTTTATTTATCTTTCTTTACA 5040
Db 5027 AAGGGAAGTCTTTGCATAGGAAAAGGGCGGCTGGCTCATTTTATTTATCTTTCTTTACA 5086
Qy 5041 CTGAATCCCAAAATCATCTTACCACAAAGGGCCAGCCTGACTGTGATTTCTCGAGTCAC 5100
Db 5087 CTGAATCCCAAAATCATCTTACCACAAAGGGCCAGCCTGACTGTGATTTCTCGAGTCAC 5146
Qy 5101 AAGAGCCATGCATCTCTGTGTTCTCACTCAGTCATGTCCCGAAGTTGTCAGATCCA 5160
Db 5147 AAGAGCCATGCATCTCTGTGTTCTCACTCAGTCATGTCCCGAAGTTGTCAGATCCA 5206
Qy 5161 GTGGCATCTGTGCTCTGCTGCACATCTTCTATTTCACCTGGCTGGCAGATCAAGTGT 5220
Db 5207 GTGGCATCTGTGCTCTGCTGCACATCTTCTATTTCACCTGGCTGGCAGATCAAGTGT 5266
Qy 5221 AACTCTGGCTTCTGGGCCAAGTTAGAAAATACCAGTCTATTTCCTTTTATTTATTTTA 5280
Db 5267 AACTCTGGCTTCTGGGCCAAGTTAGAAAATACCAGTCTATTTCCTTTTATTTATTTTA 5326
Qy 5281 TTTTATTTTATTTATGCTTTTCAGTGAGTTGATGCTTCGAAAGGCTGTGTGTTATT 5340
Db 5327 TTTTATTTTATTTATGCTTTTCAGTGAGTTGATGCTTCGAAAGGCTGTGTGTTATT 5386
Qy 5341 AGCCTTGCTGCTACTGTTGACCCCAACCCATTCCTTCCTCCCTCCCTCTTCAGC 5400
Db 5387 AGCCTTGCTGCTACTGTTGACCCCAACCCATTCCTTCCTCCCTCCCTCTTCAGC 5446
Qy 5401 CAGCCTATGATAAACAATAAGATTATTAAATGCTGGCTTCGTATCTCATTAAGACAGAGAT 5460
Db 5447 CAGCCTATGATAAACAATAAGATTATTAAATGCTGGCTTCGTATCTCATTAAGACAGAGAT 5506
Qy 5461 TGTCACTTGAATCTACTTCTATAGCATTTCAAAGTGCCACGGCCCAACACCCATGATGTT 5520
Db 5507 TGTCACTTGAATCTACTTCTATAGCATTTCAAAGTGCCACGGCCCAACACCCATGATGTT 5566
Qy 5521 CTTCAATGCTCTGAAGGTCAAGAGCCTCATTTTGTGTTTCTGGTTAGATCTCTTTCCTCC 5580
Db 5567 CTTCAATGCTCTGAAGGTCAAGAGCCTCATTTTGTGTTTCTGGTTAGATTTTTCCTCC 5626

Qy 5581 TTGCTTGAATGAATAAACCGTTTAAACAGTAGGCTCTTAGCATCACACACATAGTCAT 5640
Db 5627 TTGCTTGAATGAATAAACCGTTTAAACAGTAGGCTCTTAGCATCACACACATAGTCAT 5686
Qy 5641 TCCTCATGTTCTTGTGTTTAAACAAGCAGTTCGAGTTTCTGAGTTTAAATTAAGTCTGCAA 5700
Db 5687 TCCTCATGTTCTTGTGTTTAAACAAGCAGTTCGAGTTTCTGAGTTTAAATTAAGTCTGCAA 5744
Qy 5701 TGAGACAATTTATAAACCCATTTAGCTGGGTGGAAAATTG-TTCTCAAAAGCAATAAGTA 5759
Db 5745 TGAGACAATTTATAAACCCATTTAGCTGGGTGGAAAATTGTTCTCAAAAGCAATAAGTA 5804
Qy 5760 ATAAATCTGTTATCTGCTATTAACCTACAGTTTGAATAAGAAAGTACCAAGTCTCAGTAC 5819
Db 5805 ATAAATCTGTTATCTGCTATTAACCTACAGTTTGAATAAGAAAGTACCAAGTCTCAGTAC 5864
Qy 5820 ATTATATATCATTTGGGGTCTGAGTAACCTGGGAGTGTAGTCTTTTGAGACTTTGTAGCAC 5879
Db 5865 ATTATATATCATTTGGGGTCTGAGTAACCTGGGAGTGTAGTCTTTTGAGACTTTGTAGCAC 5924
Qy 5880 CAGTCTTTATTAGGAAAGTCTGTTGGCTTTTACAGGGCATTAGTCCCTTTGTGTTTC 5939
Db 5925 CAGTCTTTATTAGGAAAGTCTGTTGGCTTTTACAGGGCATTAGTCCCTTTGTGTTTC 5984
Qy 5940 CATGATGCTTTAAGTTCTTTTGAGTCTCATTTAAGAAATTCCTTTCTCGAAGCATGACA 5999
Db 5985 CATGATGCTTTAAGTTCTTTTGAGTCTCATTTAAGAAATTCCTTTCTCGAAGCATGACA 6044
Qy 6000 AGTGTATCGCAATCTTACATGCTCAGTCTGTTTACCTG-CGTTAGTTTGTGCTGGTTAT 6058
Db 6045 AGTGTAT-CGCAATCTTACATGCTCAGTCTGTTTACCTGTTAGTTTGTGCTGGTTAT 6103
Qy 6059 TTAATTGCACCTTTCCAGCATCATGCTTCCCTTACAAATATGATATCTTTATTTGTAC 6118
Db 6104 TTAATTGCACCTTTCCAGCATCATGCTTCCCTTACAAATATGATATCTTTATTTGTAC 6163
Qy 6119 ACTAAGGTGTGATCATGATCTGCTCCCTGTAAAGAAATTAATAAATTTTCCAG 6174
Db 6164 ACTAAGGTGTGATCATGATCTGCTCCCTGTAAAGAAATTAATAAATTTTCCAG 6219

RESULT 4
AB052621
LOCUS
DEFINITION Mus musculus DDM36E mRNA, complete cds.
ACCESSION AB052621
VERSION AB052621.1 GI:11862940
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
1. 6222
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"

AB052621
Mus musculus DDM36E mRNA, complete cds.
AB052621.1 GI:11862940
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (sites)
Murakami,H., Nakamata,T., Nakayama,T., Yamamoto,H., Hosaka,T.,
Aoyama,T., Nagayama,S., Oka,M., Kiyono,T., Sasaki,M.S., Nakamura,T.
and Toguchida,J.
Up-regulation of a ras effector and down-regulation of a cell
adhesion molecule are associated with transformation of osteoblasts
Unpublished
2 (bases 1 to 6222)
Toguchida,J., Nakamata,T., Murakami,H., Nakayama,T. and Nakamura,T.
Direct Submission
Submitted (11-DEC-2000) Junya Toguchida, Kyoto University,
Institute for Frontier Medical Sciences, 53 Kawahara-cho, Shogoin,
Sakyo-ku, Kyoto, Kyoto 606-8507, Japan
(E-mail:togjun@frontier.kyoto-u.ac.jp, Tel.81-75-751-4134,
Fax:81-75-751-4144)
Location/Qualifiers
1. 6222
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"

gene	/clone="DDM36E"		Db	470		CTCGAAGACTTCTCTCTGCAACCCCGAGTCCAGATTGTGGAGGAGAAACGGGACGACGACG		529
	/cell_line="MMC2"			481		TTTGAATGCCACACCAAGGCGCTTCCAGCCCCCATCATTTACTTTGGGAAAAGACACAGGTG		540
CDS	/note="similar to Nope (NM_020043)-vector:pDDM36BOP"		Db	530		TTTGAATGCCACACCAAGGCGCTTCCAGCCCCCATCATTTACTTTGGGAAAAGACACAGGTG		589
	1..6222			541		ACCGTGCCTGAGGAGCCCGGCTCATCTCTTCCAAAGTGGCTCTCTCCAGATCTCTAGAT		600
	/genes="DDM36E"		Db	590		ACCGTGCCTGAGGAGTCCCGGCTCATCTCTTCCAAAGTGGCTCTCTCCAGATCTCTAGAT		649
	50..3811			601		GTCCAGGACAGTGTATGACAGGCTCTTACCGCTGGTGGCCACCAATTTACAGCCCGCCCAACGA		660
	/product="DDM36E"		Db	650		GTCCAGGACAGTGTATGACAGGCTCTTACCGCTGGTGGCCACCAATTTACAGCCCGCCCAACGA		709
	/db_xref="GI:11862941"			661		TTACAGCAGGAGGCTCTCGCTCATCTGTGGCCCTCAGAGGGTCTTTTGGAGGCTTACCAAGGGG		720
	/translation="MARADTGRGLVLVLTCLLSARGELPLPQETTVKLSCDBEQLQVI		Db	710		TTACAGCAGGAGGCTCTCGCTCATCTGTGGCCCTCAGAGGGTCTTTTGGAGGCTTACCAAGGGG		769
	LGEQAVVLDCTLGATAAGPTTRVMSKGDVTLHENLHLLPNSLMLSSPLEQSDS			721		CAGGATGTGGTCTATTGTGGCAGCCCGCAGAGAAACCAACGGTAGTGTCTTGGACAGATGTA		780
	DEEARLTKWTGEGSCLASPLGVVASQVAVKLATLEDFSLPESQIVRENGAR		Db	770		CAGGATGTGGTCTATTGTGGCAGCCCGCAGAGAAACCAACGGTAGTGTCTTGGACAGATGTA		829
	FECHTKGAPLITWEEKQVTPESRLITLPLNGVLQILDVQSDAGSYRCVATNSAR			781		GTGATGGAGTGGTGGGCTCTGTCTGACCCCCACCCCTTTTGTGTCTTGGGTCCGACAGGAT		840
	QRFQBSASITVALRGSLATRGQDVVIAAPENTTVSGQSVMECVASADPTFPVSW		Qy	830		GTGATGGAGTGGTGGGCTCTGTCTGACCCCAACCCCTTTTGTGTCTTGGGTCCGACAGGAT		889
	VRQKQPISTDIVLGRNTLLIASAQPHSGVVCRAKPRTRDFATAAEILRLVLAAP			841		GGAAAGCCTATCTCCAGGATGTCTCGTTCTGGGCGGGACCAATCTACTATCTGCGCCAGC		900
	ATSOAPEALSRTRASTARFVCRASGPRPALHWDGIPLRPNRGVKGCGGSLVIT		Db	890		GGAAAGCCTATCTCCAGGATGTCTCGTTCTGGGCGGGACCAATCTACTATCTGCGCCAGC		949
	QIGLODAGYQCAENSAGTACAAAPLAVVREGLPSPATVATPLSSSVLVAVER			901		GCGCAGCTCGGCACTCTGGAGTCTATGTCTGCGGAGCCAAACAGGCCCTCTCACGGTGC		960
	PELHSQIILGSLHYOKARGVDNEYQFVAVNDTTELQVRDLPLNDTDFYVVAISQL		Db	950		GCGCAGCTCGGCACTCTGGAGTCTATGTCTGCGGAGCCAAACAGGCCCTCTCACGGTGC		1009
	GASRTSPALVHLTDVPSAAQPLTLLSPNSDIRVAMLPPLSSLSNGQVLKYIEYG			961		TTGCGCCTCGGCGCTGTGAGCTCCGAGTGTCTGTGCCCCCAGCCATCTCTGCGAGGCACCC		1020
	LGKEQVSTFTEVPGNETQTLNLSQPNKVRYSAGTGAGYVPSQMHRTPGVH		Qy	1010		TTGCGCCTCGGCGCTGTGAGCTCCGAGTGTCTGTGCCCCCAGCCATCTCTGCGAGGCGCCC		1069
	NOSHVPFAELKVRAMESLVSVMQPPPHPTQISGYKYRVEHGTREEDAGDRPPGT			1021		GAGGCGCTCTGCGGAGCGCGGCGGACGCGCGGCTCTGTGTGCGGCGGCTCTCGGGGAG		1080
	RGDOAVGVPLKXKVKQYELTOLVGRLEYEVLVAFNKHEDGYAAVWKGKTEKAPT		Db	1070		GAGGCGCTCTGCGGAGCGCGGCGGACGCGCGGCTCTGTGTGCGGCGGCTCTCGGGGAG		1129
	PDLTORGPPLPAHVHAESNSTSIWLRWKKPDPTTKIVNTVTRFGPWGLRNASLV			1081		CCACGCGCCCGGCTGCACTGTGCTGCAAGCGGATCCCGTTGCGACCCCAATGGGCGCGTC		1140
	TYTSSGEBILIGGLKPLPTTKYEPFVAVSHGVDMDGPRGVSVERSTLDRPSTPDLRL		Qy	1130		CCACGCGCCCGGCTGCACTGTGCTGCAAGCGGATCCCGTTTGGAGCCCAATGGGCGCGTC		1189
	SPUTSPRLHMCPTVEPGEIIVEYLILYNNHTPQHQWTLITLTTGNIPLSFAVHGLE			1141		AAGGTGACGGGCGGTGGCGGCGGCTGTGCTCATCTCAGATCCGCTCTGCGAGGACGCTGGC		1200
	SDTRYFKMGATTEVGPFFSRQLQVITLQTFSDSLDVHATVGTIVGVCILGLCLLA		Db	1190		AAGGTGACGGGCGGTGGCGGCGGCTGTGCTCATCTCAGATCCGCTCTGCGAGGACGCTGGC		1249
	CMCAGLRSSHREALPLSGSSGCPGNPALYTRALRPPSPVPAHLESILVHPRPDMS			1201		TACTACAGTGGTGTAGCAGAAAAACAGCGCGGAACTGCTCTGTGCGGCTGCGCCCTGGCG		1260
	PPSDVEDKAEVHLMGSSVSDCRGHSRKKISWAQGNWAGSWAGCELPQSGGPRP		Qy	1250		TACTACAGTGGTGTAGCAGAAAAACAGCGCGGAACTGCTCTGTGCGGCTGCGCCCTGGCG		1309
	ALTRALLPAGTQTLLOALVYDAIKSNGRKKPSACRNOVEARVIVHSDFGASKGC			1261		GTAGTGGTGGCGAGGCGCTGCCCAGCGCCCGCACTCGGCTCAGACCGCGCTGAGC		1320
	PDLHQLDPEEPLEATLFTSGSAGVLSQKALCPCLTVSPSPRPVSSAQVP"		Db	1310		GTAGTGGTGGCGAGGCGCTGCCCAGCGCCCGCACTCGGCTCAGACCGCGCTGAGC		1369
	BASE COUNT 1321 a 1772 c 1653 g 1476 t			1321		AGCTCTCTGTGCTGTGGAGCGGCTGTGAGTTGCAACAGGCAAAATCATTTGGC		1380
	ORIGIN		Db	1370		AGCTCTCTGTGCTGTGGAGCGGCTGTGGAGCGGCTGTGAGTTGCAACAGGCAAAATCATTTGGC		1429
				1381		TTCTCTCTTTCATCTACCAAAAGGCAAGGGAGTGGACAAATGTGGAGTTACAGTTTTCAGTA		1440
			Qy	1430		TTCTCTCTTTCATCTACCAAAAGGCAAGGGAGTGGACAAATGTGGAGTTACAGTTTTCAGTA		1489
				1441		AACAAATGACACCAACAGAGTGTGAGGTTTGGGACCTGGAAACCCAAACAGGATTTATGATTC		1500
			Db	1490		AACAAATGACACCAACAGAGTGTGAGGTTTGGGACCTGGAAACCCAAACAGGATTTATGATTC		1549
				1501		TAGGTGGTGGCTTACTTCCAGCTGGGGGCGAGCGGAACTCTCAGCCCGAGCCCTGTGTGCAT		1560
			Qy	1550		TAGGTGGTGGCTTACTTCCAGCTGGGGGCGAGCGGAACTCTCAGCCCGAGCCCTGTGTGCAT		1609


```

Qy 5937 TGCCATGATGCTTAAGTCTTTGAGTCTCAATTAAGAATCTCTTTCTCGAAGCATG 5996
Db 5985 TGCCATGATGCTTAAGTCTTTGAGTCTCAATTAAGAATCTCTTTCTCGAAGCATG 6044
Qy 5997 ACAAGTGTATGCAATACTTACATCTCACTCGTTTACTCG- GCTTAGTGTGTGCTGGGT 6055
Db 6045 ACAAGTGTAT- GCAATACTTACATCTCACTCGTTTACTCGTGTAGTGTGTGCTGGGT 6103
Qy 6056 TATTTAAATGGCACTTCCAGCATCATGCTTCTCTTCAAAATATGATATCTTTATTGT 6115
Db 6104 TATTTAAATGGCACTTCCAGCATCATGCTTCTCTTCAAAATATGATATCTTTATTGT 6163
Qy 6116 TACACTAAGTGTTCATCATGATATCTGCTCCCTGTAAGAAATTAATAAATACTATTTCCAG 6174
Db 6164 TACACTAAGTGTTCATCATGATATCTGCTCCCTGTAAGAAATTAATAAATACTATTTCCAG 6222

RESULT 5
AK122535
LOCUS AK122535 6301 bp mRNA linear ROD 15-MAR-2003
DEFINITION Mus musculus mRNA for mKIAA1628 protein.
ACCESSION AK122535
VERSION AK122535.1 GI:28972801
KEYWORDS FLI_CDNA.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Aizawa,H., Yuasa,S.,
Nakajima,D., Nagase,T., Ohara,O. and Koga,H.
TITLE Prediction of the coding sequences of mouse homologues of KIAA
gene: II. The complete nucleotide sequences of 400 mouse
KIAA-homologous cDNAs identified by screening of terminal sequences
of cDNA clones randomly sampled from size-fractionated libraries
DNA Res. 10, 35-48 (2003)
JOURNAL
REFERENCE 2 (bases 1 to 6301)
AUTHORS Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.
TITLE Direct Submission
JOURNAL Submitted (07-FEB-2003) Hiasahi Koga, Kazusa DNA Research
Institute, Laboratory for Genome Informatics; 2-6-7
Kazusa-kamatori, Kisarazu, Chiba 292-0818, Japan
(E-mail:mous@kazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)
COMMENT The CREATE program supported by Japan science and technology
corporation; cDNA full insert sequencing: Kazusa DNA Research
Institute; cDNA library construction, clone selection and 5'- &
3'-end one pass sequencing.
FEATURES
source
1..6301
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="mbg00013"
/tissue_type="brain"
/dev_stage="adult"
/note="vector:modified pBC SK+"
gene 1..6301
/gene="mKIAA1628"
CDS <1..3887
/gene="mKIAA1628"
/note="CDS is predicted by in silico analysis. Start codon
is not identified."
/evidence=not_experimental
/product="mKIAA1628 protein"
/protein_id="BAC65817.1"
/db_xref="GI:28972802"
/translation="EGSQNPNSAMHGRGELPLPQRTTVTKLSCDRGLQVILGPQAVV
LDCTLGATGAPPTVMTSKDQDTVLHENHLLPNSLWLSPLEQSDSDDEALRI
WKVTGYSCLAHSPGLGVASQAVKLTLEDLSLPESQIVBENGTAPECHTKGL
PAPITFKDQVTPESRLITLPLNGVLQILDQSDAGSYRCVATNSARQFSQAS
LTVALRSLGTEATRGQDVVIAAPENTTVVSQGSVMBCVASADPTPFVSVVRQDQSK

```

```

KELGIQTSGCDAALQCSGNLTVGEVGEVTPASPSSLLLLCHQMESLSRMSFWAG
PIYSSPARSLGTLESMAEPTSPARVISLURLISSBCLLPQRRRRRRRGRGPAIR
ASCACAPGSPRECTCTGSGDRPMGRVKVGGGSLVTQIGLADGAYQCVAEHS
AGTACAAAPLAVVVRBGLPBPATRTATPLSSSSVIVAMERPELHSQIIGFSLHYK
ARGVDNVYQFVANNDDTLEQVDRDLBENTDYFVVAISYOLGASRSTSPALVHTLDIV
PSAAPOLJTLSSPNPDIIRVAVLPLPSSLSNGQVLYKIEYGLKBEQDOVPSVGRNE
TQILTSLQPNKYVRVIRISAGTCAGYQVPSQWQHRTPGVHNSHQVFPAPAEALKYRAK
MESLIVSWQPPPHPTQISGYKLTYRBEVGTBERADGPRPGGRGDQAWDVGPVRLKKV
KQYELTQLVPRGLYEVKLVAFNKHEDEGYAAVMKGTESKAPTDPLOIGRPPLPAHVH
ABENSSTSIWLWKKCPDFTTKIVNTYTRFPGMGLRNASLVTYTTSSGDIILIGLHP
FTKYEPVAVQSHGMDQPGFVSVRSSTLPDRPSTPPSDRLSLVPLTPTSLHMCPTTB
PNEIEVYLJLYSNHNTQPEHOMTLLTTEGNIISAEVHGLESDTRVFPKMGARTVEUP
GPPSRLEQVLTLOKTFSDSLDVHATGIIIVGVLGLLCLLACWAGLRSHREALPG
LSSSGTGNPALYTRAKRLSPVPAHELSLHVPRPDQSPSPSPVEDKAEVHSLMG
GVSVDCHGSHSKRIKISWAGQPNWAGWAGCELPQSGSPRALTRALLPAGTQTLL
LOALVYDAIKSNRKKFSPACRNQVEAEVIVHSDFGASKGCPDILHLDLPEPLTAEB
TLPTSGAVDLQSGADWLRELGGCQPTTSGPERLTCLPEAASACSCSDLPQSTAE
EAPKSCOPKALCPLTVPSPLPRAPVSSAQP"
BASE COUNT 1362 a 1770 c 1669 g 1500 t
ORIGIN
Query Match 94.1%; Score 5810.2; DB 10; Length 6301;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 6077; Conservative 0; Mismatches 33; Indels 157; Gaps 9;
Qy 62 GCGGGAGCTGCCATTGCCCCAGGAGACAATGTCAAGCTGAGCTGTGATGAGGACCCC 121
Db 40 GAGGGAGCTGCCATTGCCCCAGGAGACAATGTCAAGCTGAGCTGTGATGAGGACCCC 99
Qy 122 TGCAGTGTATCTGGGCGCTGAGCAGGCTGTGGTCTGGACTGACATTTGGGGCTACAG 181
Db 100 TGCAGTGTATCTGGGCGCTGAGCAGGCTGTGGTCTGGACTGACATTTGGGGCTACAG 159
Qy 182 CTGCTGGGCGCTCCAGCAGGAGTGCATGTGAGCAGAGTGGAGACACTGTACTAGAGCATG 241
Db 160 CTGCTGGGCGCTCCAGCAGGAGTGCATGTGAGCAGAGTGGAGACACTGTACTAGAGCATG 219
Qy 242 AGAACCTGCACTGCTACCCAAATGGCTCCCTGTGGTGTCTCTACCCCTAGACAGAAG 301
Db 220 AGAACCTGCACTGCTACCCAAATGGCTCCCTGTGGTGTCTCTACCCCTAGACAGAAG 279
Qy 302 ACAGCATGTAGAGAAAGCTCTTAGGATCTGGAAGTCACTGAGGGCAGCTATTCTGTGTC 361
Db 280 ACAGCATGTAGAGAAAGCTCTTAGGATCTGGAAGTCACTGAGGGCAGCTATTCTGTGTC 339
Qy 362 TGGCCCAAGCCGCTAGGAGTGTGTCAGCAGGCTGTGTGTGCTCAAGCTTCCCAAC 421
Db 340 TGGCCCAAGCCGCTAGGAGTGTGTCAGCAGGCTGTGTGTGCTCAAGCTTCCCAAC 399
Qy 422 TCGAAGACTTCTCTCTGACCCCGAGTCCAGAGTTGTGGAGGAGAACGGGACAGCAGCT 481
Db 400 TCGAAGACTTCTCTCTGACCCCGAGTCCAGAGTTGTGGAGGAGAACGGGACAGCAGCT 459
Qy 482 TTGAATGCCACACCAAGGGCTTCCAGCCGCCATCATTAATTGAGGAAAAGACAGGTGA 541
Db 460 TTGAATGCCACACCAAGGGCTTCCAGCCGCCATCATTAATTGAGGAAAAGACAGGTGA 519
Qy 542 CCGTGCCTGAGGAGCCCGGCTCATCACTCTTCCCAAGTGGCTCCTCCAGATCTTAGATG 601
Db 520 CCGTGCCTGAGGAGTCCCGGCTCATCACTCTTCCCAAGTGGCTCCTCCAGATCTTAGATG 579
Qy 602 TCCAGGACAGTGTAGAGGCTCTTACCGCTGCGTGGCCACCAATTCAGCCCGCCCAAGCAT 661
Db 580 TCCAGGACAGTGTAGAGGCTCTTACCGCTGCGTGGCCACCAATTCAGCCCGCCCAAGCAT 639
Qy 662 TCAGCCAGGAGGCTCCCTCATCTGTGGCCCTCAGAGGGTCTTTGAGAGGCTTACAGGGGGC 721
Db 640 TCAGCCAGGAGGCTCCCTCATCTGTGGCCCTCAGAGGGTCTTTGAGAGGCTTACAGGGGGC 699
Qy 722 AGGATGTGTCATTTGGGAGCCGCCAGACACCAAGTGTCTTGGAGAGTGTCTGGAGATGTAG 781
Db 700 AGGATGTGTCATTTGGGAGCCGCCAGACACCAAGTGTCTTGGAGAGTGTCTGGAGATGTAG 759

```

QY 782 TGATGAGTGCCTGCTGAGCCACCCCTTTTGTGCTCTGGTCCGACAG---- 837
DB 760 TGATGAGTGCCTGCTGAGCCACCCCTTTTGTGCTCTGGTCCGACAGGGG 819
QY 838 ----- 837
DB 820 ATCAGAGTAAGAGAACTGGGATTAACACAGCGCTGTGATGAGCTTTGACAGGAT 879
QY 838 ----- 837
DB 880 GTAGTGGGAACCTCGTAGTTGGAGAGGGGTGAGAGGGACAGTTTCAGCGTCCCTCC 939
QY 838 -----GATGGAAAGCCTATCTCCACCGATGTCTATCTGCTGCG 875
DB 940 GCCTGCTCTCTCTGTCACCAAGTGAAGACCTATCTCCACGATGTCTATCTGCG 999
QY 876 CCGGACCAATCTATCATGCGCCAGCGCGAGCTCTGCGACCTCTGGAGTCTATCTGCG 935
DB 1000 CCGGACCAATCTATCATGCGCCAGCGCGAGCTCTGCGACCTCTGGAGTCTATCTGCG 1059
QY 936 AGCCACAAGCCCTCAGCGTGAATCTGCGCACTGCGGCTGCTGAGCTCCGAGTCTGC 995
DB 1060 AGCCACAAGCCCGCAGCGTGAATCTGCGCACTGCGGCTGCTGAGCTCCGAGTCTGC 1119
QY 996 TGCCCCAGCCATCTCGCAGGACACCGAGGCGCTCTCGCGACCGCGGCCAGCACCGCG 1055
DB 1120 TGCCCCAGCCATCTCGCAGGCGCGCGAGGCGCTCTCGCGACCGCGGCCAGCACCGCG 1179
QY 1056 CTTGCTGTGCGGCGCTCGGAGAGCAACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 1115
DB 1180 CTTGCTGTGCGGCGCTCGGAGAGCAACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 1239
QY 1116 CCGTTCGACCAATGGG--CGGTCGAAGTGCAGGCGCTGCGCGAGCTCGGTCATCA 1174
DB 1240 CCGTTCGACCAATGGGCGCGCTCAAGTGCAGGCGCTGCGCGAGCTCGGTCATCA 1299
QY 1175 CTCAGATCGGCTCGAGGACGCTGGCTACTACAGTGCCTGAGCAAGAAACAGCGCGGAA 1234
DB 1300 CTCAGATCGGCTTCAGGACGCTGGCTACTACCGTGCCTGAGCAAGAAACAGCGCGGAA 1359
QY 1235 CTGCTGTGCGGCTCGCGCTCGGAGTGTGCTGCGAGGCGCTGCGCGCGCGG 1294
DB 1360 CTGCTGTGCGGCTCGCGCTCGGAGTGTGCTGCGAGGCGCTGCGCGCGCGG 1419
QY 1295 CTGCGGTTCAGGCAACCGCGCTGAGCAGCTCTCTGCTGCTGCTGCTGCTGCTGCTG 1354
DB 1420 CTGCGGTTCAGGCAACCGCGCTGAGCAGCTCTCTGCTGCTGCTGCTGCTGCTGCTG 1479
QY 1355 AGTTGCAAGGCAAAATCAATGGCTTCTCTCTTCACTACCAAAAGGCAAGGAGTGG 1414
DB 1480 AGTTGCAAGGCAAAATCAATGGCTTCTCTCTTCACTACCAAAAGGCAAGGAGTGG 1539
QY 1415 ACAATGGAGTACCAAGTTTCAGTAACAAATGACACACAGAGCTGCGAGTTGCGGAC 1474
DB 1540 ACAATGGAGTACCAAGTTTCAGTAACAAATGACACACAGAGCTGCGAGTTGCGGAC 1599
QY 1475 TGGAAACCAACAGGATTAAGTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1534
DB 1600 TGGAAACCAACAGGATTAAGTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1659
QY 1535 GAACCTCCAGCCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1594
DB 1660 GAACCTCCAGCCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1719
QY 1595 TTACCTTTGTCAGCCCAACCCCTCGGACATCAGGCTGGGATGCTGCTGCTGCTGCTGCTG 1654
DB 1720 TTACCTTTGTCAGCCCAACCCCTCGGACATCAGGCTGGGATGCTGCTGCTGCTGCTGCTG 1779
QY 1655 GCCTGAGCAATGGAAGTCTGAGTCAAGTACAGTACAGTACAGTACAGTACAGTACAG 1711
DB 1780 GCCTGAGCAATGGAAGTCTGAGTCAAGTACAGTACAGTACAGTACAGTACAGTACAG 1839
QY 1712 ATCAGGTTTTCTCCACCGAGTGCCTTGGAAATGAGACACAACCTTACGTTAACTCACTTC 1771

DB 1840 ATCAGGTTTTCTCCACCGAGTGCCTTGGAAATGAGACACAACCTTAACTCACTTC 1899
QY 1772 AGCCAAAACAAGTGTACCGATCCCGATTTTCACTGCGCCTGCGCTGCTGCTGCTGCTG 1831
DB 1900 AGCCAAAACAAGTGTACCGATTTTCACTGCGCCTGCGCTGCTGCTGCTGCTGCTG 1959
QY 1832 CTTTCTCAGTGGATGACGACAGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1891
DB 1960 CTTTCTCAGTGGATGACGACAGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2019
QY 1892 CCGCTTCAGAAATTGAAGTGTAGGCGAAAGATGGAATGCTTCTGCTGCTGCTGCTGCTG 1951
DB 2020 CCGCTTCAGAAATTGAAGTGTAGGCGAAAGATGGAATGCTTCTGCTGCTGCTGCTGCTG 2079
QY 1952 CCGCTTCACCCACCCAGATCTCTGATCAAACTCTTCTGATCAAACTCTTCTGATCAAACT 2011
DB 2080 CCGCTTCACCCACCCAGATCTCTGATCAAACTCTTCTGATCAAACTCTTCTGATCAAACT 2139
QY 2012 AGGAGGCGAGTGGTGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2071
DB 2140 AGGAGGCGAGTGGTGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2199
QY 2072 CCGTGGCTGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2131
DB 2200 CCGTGGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2259
QY 2132 CGTACGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2191
DB 2260 TGTACGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2319
QY 2192 AGGGAAGACGAGGAGGCG 2251
DB 2320 AGGGAAGACGAGGAGGCG 2379
QY 2252 CTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2311
DB 2380 CTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2439
QY 2312 AGCCAGACTTTTACCACTGCTCAAGATTTGTCAACTTACTGCTGCTGCTGCTGCTGCTG 2371
DB 2440 AGCCAGACTTTTACCACTGCTCAAGATTTGTCAACTTACTGCTGCTGCTGCTGCTGCTG 2499
QY 2372 TCAGGAATCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2431
DB 2500 TCAGGAATCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2559
QY 2432 GCCTGAAACCAATTTTACCAAGTACGAGTTTGGTACAGTCCCAAGGAGTGAATGAGATG 2491
DB 2560 GCCTGAAACCAATTTTACCAAGTACGAGTTTGGTACAGTCCCAAGGAGTGAATGAGATG 2619
QY 2492 GCGCTTTTGGCTCGCTGCTGAGAACGCTCCACCTGCTGAGACCGGCTTCAACACCTCTT 2551
DB 2620 GCGCTTTTGGCTCGCTGCTGAGAACGCTCCACCTGCTGAGACCGGCTTCAACACCTCTT 2679
QY 2552 CTGACCTGCGCTGAGCCCTGACACCATCAACCGTTCGCTTACACTGCTGCTGCTGCTGCTG 2611
DB 2680 CTGACCTGCGCTGAGCCCTGACACCATCAACCGTTCGCTTACACTGCTGCTGCTGCTGCTG 2739
QY 2612 CCGAGCCCAATGGTGAAGTGTGGAGTATCTAAATCTCTACAGCAACCAACCAACCCAGC 2671
DB 2740 CCGAGCCCAATGGTGAAGTGTGGAGTATCTAAATCTCTACAGCAACCAACCAACCCAGC 2799
QY 2672 CCGAACACAGTGGACACTGCTCACCACAGAGGAAACATCTTCACTGAGGAGTCCATG 2731
DB 2800 CCGAACACAGTGGACACTGCTCACCACAGAGGAAACATCTTCACTGAGGAGTCCATG 2859
QY 2732 GCCTAGAGTGAACACTCGGTATTTCTTCAAGATGGAGCCCGCAGAGGTTGGGCGCTG 2791
DB 2860 GCCTAGAGTGAACACTCGGTATTTCTTCAAGATGGAGCCCGCAGAGGTTGGGCGCTG 2919
QY 2792 GCGCTTTTCCCGCTTGGAGTGTGATTACTCTGCAAGAGACATTTCTCAGACTCTCTGG 2851

[illegible]

Qy	5012	TGCTCATTTATTTATCTTTCTTTACACTGAATCCCAAAATCATCTTTACCAAAAGG	5071
Db	5138	TGGCTCATTTTATTTTATCTTTCTTTACACTGAATCCCAAAATCATCTTTACCAAAAGG	5197
Qy	5072	CCAAGCTGACTGGTATTTCTCGAGTCACAAGAGCCATGCACTCTCTGTGTTCTCAC	5131
Db	5198	CCAAGCTGACTGGTATTTCTCGAGTCACAAGAGCCATGCACTCTCTGTGTTCTCAC	5257
Qy	5132	TCAGTCATGTCACAGAAATGTTCAGATCCAGTGGCATCTGTGCTCTTGTGCTGCA	5191
Db	5258	TCAGTCATGTCACAGAAATGTTCAGATCCAGTGGCATCTGTGCTCTTGTGCTGCA	5317
Qy	5192	TATTTCAACTGGCTGGCAATCAAGTGTTAACTCTGGCTTCGGGCAAGTTAGAAATAA	5251
Db	5318	TATTTCAACTGGCTGGCAATCAAGTGTTAACTCTGGCTTCGGGCAAGTTAGAAATAA	5377
Qy	5252	CCAGTCATATTTTCCCTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	5311
Db	5378	CCAGTCATATTTTCCCTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	5437
Qy	5312	TGTAGCTTCTGAAAGCGTCTGTGTTTATTAGCCTTGTGTGCTCACTCATGTTTGACCCAC	5371
Db	5438	TGTAGCTTCTGAAAGCGTCTGTGTTTATTAGCCTTGTGTGCTCACTCATGTTTGACCCAC	5497
Qy	5372	CCACATTTCTTCTCTCCCTCTTCAGCCAGGCTATGATAACACATAAAGATTTTAAATG	5431
Db	5498	CCACATTTCTTCTCTCCCTCTTCAGCCAGGCTATGATAACACATAAAGATTTTAAATG	5557
Qy	5432	CTGGCTTCGTATCTCATTAAGACAGGATGTGCATTTGAACACTCTTATAGCATTTCAA	5491
Db	5558	CTGGCTTCGTATCTCATTAAGACAGGATGTGCATTTGAACACTCTTATAGCATTTCAA	5617
Qy	5492	GTGGCCACGGCCAAACACCACCGTATGTTCTTCATTTGCTCTGAAGGTCGAAGGCTCAT	5551
Db	5618	GTGGCCACGGCCAAACACCACCGTATGTTCTTCATTTGCTCTGAAGGTCGAAGGCTCAT	5677
Qy	5552	TTGTTTTCCTGGTTAGATCTTTTCTCTCTGCTTGAATGAAATAACCGTTTAAACAGT	5611
Db	5678	TTGTTTTCCTGGTTAGATCTTTTCTCTCTGCTTGAATGAAATAACCGTTTAAACAGT	5737
Qy	5612	AGGCTCTTAGCATCACACCATAGTCATCTCATGTTCTGCTTTTAAACAGCACATTGAG	5671
Db	5738	AGGCTCTTAGCATCACACCATAGTCATCTCATGTTCTGCTTTTAAACAGCACATTGAG	5795
Qy	5672	GTTCTGGGTTTAAATTAATAGCTGCAATAGACAAATTTATAACCATTTAGGCTGGGTG	5731
Db	5796	GTTCTGGGTTTAAATTAATAGCTGCAATAGACAAATTTATAACCATTTAGGCTGGGTG	5855
Qy	5732	GAATAATG-TTCTCAAAGCAATAAGTATAATACTGGTATCTGCTATAACTCAAGT	5790
Db	5856	GAATAATGTTTCTCAAAGCAATAAGTATAATACTGGTATCTGCTATAACTCAAGT	5915
Qy	5791	TGATAGAAGTAGCCAGAACTCACTAGCATTTATATGATGGGTTCTGAGTAACTGG	5850
Db	5916	TGATAGAAGTAGCCAGAACTCACTAGCATTTATATGATTTGGGTTCTGAGTAACTGG	5975
Qy	5851	GGAGTGTTAGCTTTGTGACCTTTGTAGCACACAGTCTTATTAGGAAAGTCTGTTGGCTTT	5910
Db	5976	GGAGTGTTAGCTTTGTGACCTTTGTAGCACACAGTCTTATTAGGAAAGTCTGTTGGCTTT	6035
Qy	5911	TACAGGSCATTAGTCCCTTTGTGCTTTGCCATGGATGCCCTTAAAGTTCTTTGGAGTCTCAT	5970
Db	6036	TACAGGSCATTAGTCCCTTTGTGCTTTGCCATGGATGCCCTTAAAGTTCTTTGGAGTCTCAT	6095
Qy	5971	TTAAGAAATTCCTTTCTCGAAGCATGACAAGTGTATCGCAATTAATTACATGCTCACTCGT	6030
Db	6096	TTAAGAAATTCCTTTCTCGAAGCATGACAAGTGTATCGCAATTAATTACATGCTCACTCGT	6154
Qy	6031	TTACCTG-GCTTAGTTTGTGCTGGTTATTTTAAATTCGACCTTCCAGCATCATGCTTCCCTC	6089
Db	6155	TTACCTGTGCTTAGTTTGTGCTGGTTATTTTAAATTTGCACCTTCCAGCATCATGCTTCCCTC	6214
Qy	6090	CTTACAAATATGATATTTCTTTATTTGTATACACTAAGGTTGTCATCATGATCTCTGCTCCCTGT	6149

Db	6215	CTTACAAATATGATATCTTTATTTGTACACTAAGGTGTTGATCATGATATCTGTCCCTGT	6274
Qy	6150	AAAGAATTAATAAACTATTTTCCAGAC	6176
Db	6275	AAAGATTAATAAACTATTTTCCAGAC	6301
RESULT 6			
AX191303			
LOCUS	AX191303	3756 bp	DNA linear PAT 15-AUG-2001
DEFINITION	Sequence 45 from Patent WO0149714.		
ACCESSION	AX191303		
VERSION	AX191303.1	GI:15209554	
KEYWORDS			
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS	Salbaum,J.M.		
TITLE	Nope polypeptides, encoding nucleic acids and methods of use		
JOURNAL	Patent: WO 0149714-A 45 12-JUL-2001;		
FEATURES	Neurosciences Research Foundation Inc. (US)		
source	Location/Qualifiers		
	1..3756		
	/organism="Mus musculus"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:10090"		
BASE COUNT	774 a 1167 c 1105 g 710 t		
ORIGIN			
Query Match	60.8%;	Score 3756;	DB 6; Length 3756;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 3756;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	ATGCGCGGCGGACACAGCGCGCGGCTCTGTGTCTGACCTTCTGCCTGCTGTCTCGCG	60
Db	1	ATGCGCGGCGGACACAGCGCGCGGCTCTGTGTCTGACCTTCTGCCTGCTGTCTCGCG	60
Qy	61	CGCGGGGAGCTGCCATTGCCCCAGGAGACAACTGTCAAGCTGAGCTGTGATGAGGGACCC	120
Db	61	CGCGGGGAGCTGCCATTGCCCCAGGAGACAACTGTCAAGCTGAGCTGTGATGAGGGACCC	120
Qy	121	CTGCAAGTGATCTCTGGGCCCTTGAGCAGGCTGTGGTGTCTGGACTGTGCACCTTTGGGGGCTACA	180
Db	121	CTGCAAGTGATCTCTGGGCCCTTGAGCAGGCTGTGGTGTCTGGACTGTGCACCTTTGGGGGCTACA	180
Qy	181	GCTGCTGGGCTCGACAGGGTGACATGAGCAAGGATGAGACACTGTACTAGAGCAT	240
Db	181	GCTGCTGGGCTCGACAGGGTGACATGAGCAAGGATGAGACACTGTACTAGAGCAT	240
Qy	241	GAGAACCTGCACCTGCTACCAATGGCTCCTCTGTGGCTGTCTCACCCCTAGACAAAGAA	300
Db	241	GAGAACCTGCACCTGCTACCAATGGCTCCTCTGTGGCTGTCTCACCCCTAGACAAAGAA	300
Qy	301	GACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTCTGT	360
Db	301	GACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTCTGT	360
Qy	361	CTGCCCCACAGCCCGCTAGAGTGGTGGCCAGCCAGGTTGCTGTGGTCAAGCTTGCCACA	420
Db	361	CTGCCCCACAGCCCGCTAGAGTGGTGGCCAGCCAGGTTGCTGTGGTCAAGCTTGCCACA	420
Qy	421	CTCGAAGACTTCTCTGTGCAACCCGAGTCCAGATTCTGGAGGAGAACGGGACAGCAGC	480
Db	421	CTCGAAGACTTCTCTGTGCAACCCGAGTCCAGATTCTGGAGGAGAACGGGACAGCAGC	480
Qy	481	TTTGAATGCCACACAAAGGGGCTTCCAGCCCCCATCATTTACTTGGGAAAGGACACAGTG	540
Db	481	TTTGAATGCCACACAAAGGGGCTTCCAGCCCCCATCATTTACTTGGGAAAGGACACAGTG	540
Qy	541	ACCGTGCTGAGGAGCCCGGCTCATCTCTTCCCAAGTGGCTCTCCAGATCCTTAGAT	600

|||||
541 ACCGTGCTGAGGAGCCCGGCTCATCACTCTTCCCAAGTGCTCTCCAGATCCTAGAT 600 Db
|||||
601 GTCCAGGACAGTGTAGTGCAGGCTCTTACCGCTGCGTGGCCACCAATTCAGCCCGCCAAACGA 660 Qy
|||||
601 GTCCAGGACAGTGTAGTGCAGGCTCTTACCGCTGCGTGGCCACCAATTCAGCCCGCCAAACGA 660 Db
|||||
661 TTCAAGCCAGGAGGCTCGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTTACCAAGGGG 720 Qy
|||||
661 TTCAAGCCAGGAGGCTCGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTTACCAAGGGG 720 Db
|||||
721 CAGATGTGGTCACTTGTGGAGCCCGCCAGAGAACCAACGCTAGTGTCTGGACAGATGTA 780 Qy
|||||
721 CAGATGTGGTCACTTGTGGAGCCCGCCAGAGAACCAACGCTAGTGTCTGGACAGATGTA 780 Db
|||||
781 GTGATGGAGTGGTGGCTCTGCTGACCCCAACCCCTTTTGTGTCTTGGTCCGACAGAT 840 Qy
|||||
841 GGAAAGCCTATCTCCACGATGTCAATCGTTCTGGGCGGACCAATCTACTCATCGCCAGC 900 Db
|||||
841 GGAAAGCCTATCTCCACGATGTCAATCGTTCTGGGCGGACCAATCTACTCATCGCCAGC 900 Db
|||||
901 GCGCAGCTCGGCACTCTGAGTCTATGTCTGCGGAGCCACAGCCCTCACCGGTGAC 960 Qy
|||||
901 GCGCAGCTCGGCACTCTGAGTCTATGTCTGCGGAGCCACAGCCCTCACCGGTGAC 960 Db
|||||
961 TTCCGCCACTGCGGCTGCTGAGCTCCGAGTGTCTGCGCCCGACGCACTCTCGCAGGCAACC 1020 Qy
|||||
961 TTCCGCCACTGCGGCTGCTGAGCTCCGAGTGTCTGCGCCCGACGCACTCTCGCAGGCAACC 1020 Db
|||||
1021 GAGCGCTCTCGCGGAGCGCGGCGCAGCACCGCGGCTGTGTGTGCGCGGCGTCCGGGAG 1080 Qy
|||||
1021 GAGCGCTCTCGCGGAGCGCGGCGCAGCACCGCGGCTGTGTGTGCGCGGCGTCCGGGAG 1080 Db
|||||
1081 CCAGGCGCGGCTGCACTGGCTGCAAGCGGATCCGTTGCGACCAATGGGCGGTC 1140 Qy
|||||
1081 CCAGGCGCGGCTGCACTGGCTGCAAGCGGATCCGTTGCGACCAATGGGCGGTC 1140 Db
|||||
1141 AAGGTGACGGCGGTGGCGGAGCTTGGTCACTCATCTCAGATCGGCTGCGAGGAGCTGGC 1200 Qy
|||||
1141 AAGGTGACGGCGGTGGCGGAGCTTGGTCACTCATCTCAGATCGGCTGCGAGGAGCTGGC 1200 Db
|||||
1201 TACTACCAAGTGTGAGCAAAAACAGCGCGGAACTGCTGTGCGGCTGCGCCCTGGCG 1260 Qy
|||||
1201 TACTACCAAGTGTGAGCAAAAACAGCGCGGAACTGCTGTGCGGCTGCGCCCTGGCG 1260 Db
|||||
1261 GTAGTGTGCGGAGGGGCTGCCAGCGCCCGGACTGGGTTCACAGCCCGGCTGAGC 1320 Qy
|||||
1261 GTAGTGTGCGGAGGGGCTGCCAGCGCCCGGACTGGGTTCACAGCCCGGCTGAGC 1320 Db
|||||
1321 AGCTCTCTGTGCTGGTGGCTGGAGCGGCTGAGTTGACAGCGAGCAATCAATTGGC 1380 Qy
|||||
1321 AGCTCTCTGTGCTGGTGGCTGGAGCGGCTGAGTTGACAGCGAGCAATCAATTGGC 1380 Db
|||||
1381 TTCTCTCTTCACTACAAAAGGCAAGGGAGTGGACAAATGTGGAGTACCAAGTTTGCAGTA 1440 Qy
|||||
1381 TTCTCTCTTCACTACAAAAGGCAAGGGAGTGGACAAATGTGGAGTACCAAGTTTGCAGTA 1440 Db
|||||
1441 AACATGACACCAAGAGTGTGAGGTTGGAGCTTGGACCTGAGTTGCAAGCGAGCAATCAATTGGC 1500 Qy
|||||
1441 AACATGACACCAAGAGTGTGAGGTTGGAGCTTGGAGCTTGGAGCTTGGAGTTGAGTTTC 1500 Db
|||||
1501 TACGTGTGGCTACTCCAGCTGGGGCCAGCGAACCTCCAGCCAGCCCTGCTGTGAT 1560 Qy
|||||
1501 TACGTGTGGCTACTCCAGCTGGGGCCAGCGAACCTCCAGCCAGCCCTGCTGTGAT 1560 Db
|||||
1561 ACATGGAAGATGTCCCCAGCGCAGCACCCAGCTTACTTGTCCAGCCCAACCCCTCG 1620 Qy
|||||
1561 ACATGGAAGATGTCCCCAGCGCAGCACCCAGCTTACTTGTCCAGCCCAACCCCTCG 1620 Db
|||||
1621 GACATCAGGGTGGATGGCTGCCCTGCCCTCAGCTGAGCAATGGACAGGTCTGAAG 1680 Qy
|||||

1621 GACATCAGGGTGGATGGCTGCCCTGCCCTCAGCTGAGCAATGGACAGGTGCTGAAG 1680 Db
|||||
1681 TACAAGATACAGTACCGTTTTGGGAAAGAAAGATCAGGTTTTCTCCACCGAGGTGCTTGGGA 1740 Qy
|||||
1681 TACAAGATACAGTACCGTTTTGGGAAAGAAAGATCAGGTTTTCTCCACCGAGGTGCTTGGGA 1740 Db
|||||
1741 AATGAGACACAACTTACGTTAAACTCACTTCAGCCAAACAAAGTGTACCGAGTCCGATTT 1800 Qy
|||||
1741 AATGAGACACAACTTACGTTAAACTCACTTCAGCCAAACAAAGTGTACCGAGTCCGATTT 1800 Db
|||||
1801 TCAGCTGGCACTGGCGCTGCTATGAGTCCCTTCTCAGTGGATGACAGCACAGACACCT 1860 Qy
|||||
1801 TCAGCTGGCACTGGCGCTGCTATGAGTCCCTTCTCAGTGGATGACAGCACAGACACCT 1860 Db
|||||
1861 GGTGTGCACAAACAGAGCCATGTTCCTTTGCCCTGCGAGAAATGAAAGGTGAGGGCAAAG 1920 Qy
|||||
1861 GGTGTGCACAAACAGAGCCATGTTCCTTTGCCCTGCGAGAAATGAAAGGTGAGGGCAAAG 1920 Db
|||||
1921 ATGGAGTCCCTGGTGGTGTATGGCAGCCGCGCCCTCACCCACCCAGATCTCTGGATAC 1980 Qy
|||||
1921 ATGGAGTCCCTGGTGGTGTATGGCAGCCGCGCCCTCACCCACCCAGATCTCTGGATAC 1980 Db
|||||
1981 AAATCTTACTGGGAGAGGTGGAAACAGAGAGGAGGACAGATGCTGACCCGCCCCAGGG 2040 Qy
|||||
1981 AAATCTTACTGGGAGAGGTGGAAACAGAGAGGAGGACAGATGCTGACCCGCCCCAGGG 2040 Db
|||||
2041 GGTGTGGAGATCAAGCTTGGGACGTGCGGCCGTGCGGTGAGAAAGAAAGTGAAGCAG 2100 Qy
|||||
2041 GGTGTGGAGATCAAGCTTGGGACGTGCGGCCGTGCGGTGAGAAAGAAAGTGAAGCAG 2100 Db
|||||
2101 TATGAACTGACCCAGTTAGTCCCTGGCAGCGCGTACGAGTGAAGTCTGTAGCTTTCAAC 2160 Qy
|||||
2101 TATGAACTGACCCAGTTAGTCCCTGGCAGCGCGTACGAGTGAAGTCTGTAGCTTTCAAC 2160 Db
|||||
2161 AAACACAGAGACGCGTACGCTGTGTGAAAGGCAAGACGAGAGGGCGCCACCGCA 2220 Qy
|||||
2161 AAACACAGAGACGCGTACGCTGTGTGAAAGGCAAGACGAGAGGGCGCCACCGCA 2220 Db
|||||
2221 GACCTGCTATCCAGAGGGGGCCACCGCTGCTCTGCCCCATGTCTCCAGCAGAGTCAAAAC 2280 Qy
|||||
2221 GACCTGCTATCCAGAGGGGGCCACCGCTGCTCTGCCCCATGTCTCCAGCAGAGTCAAAAC 2280 Db
|||||
2281 AGCTCCACTTCCATTTGGCTTGGTGAAGAGCAGACTTTACCACTGTCAAGATTGTC 2340 Qy
|||||
2281 AGCTCCACTTCCATTTGGCTTGGTGAAGAGCAGACTTTACCACTGTCAAGATTGTC 2340 Db
|||||
2341 AACTACACTGTACGCTTGGCGCCCTGGGGGCTCAGGAAATGCTTCCCTGGTCACTTAT 2400 Qy
|||||
2341 AACTACACTGTACGCTTGGCGCCCTGGGGGCTCAGGAAATGCTTCCCTGGTCACTTAT 2400 Db
|||||
2401 ACCAGCTCTGGAGAGACATTTCTCATTTGGCGGCTGAAACCAATTTACCAAGTACGAGTTT 2460 Qy
|||||
2401 ACCAGCTCTGGAGAGACATTTCTCATTTGGCGGCTGAAACCAATTTACCAAGTACGAGTTT 2460 Db
|||||
2461 GCGGTACAGTCCCAAGAGTGGATATGGATGGGCGCTTTGGCTCCGTGTAGAAACGCTTC 2520 Qy
|||||
2461 GCGGTACAGTCCCAAGAGTGGATATGGATGGGCGCTTTGGCTCCGTGTAGAAACGCTTC 2520 Db
|||||
2521 ACCCTGCTGACCGGCTTCAACACCTCTTCTGACCTGGCGCTGAGCGCCCTGACACCA 2580 Qy
|||||
2521 ACCCTGCTGACCGGCTTCAACACCTCTTCTGACCTGGCGCTGAGCGCCCTGACACCA 2580 Db
|||||
2581 TCCACCTTGGTGTACACTGCTGTGCCCCCAAGAGGCCAATGCTGAGATTGTGAGTAT 2640 Qy
|||||
2581 TCCACCTTGGTGTACACTGCTGTGCCCCCAAGAGGCCAATGCTGAGATTGTGAGTAT 2640 Db
|||||
2641 CTAATTTCTTACAGCAACAAACCAACCGCCGAAACCAAGTGGACACTGCTCACCA 2700 Qy
|||||
2641 CTAATTTCTTACAGCAACAAACCAACCGCCGAAACCAAGTGGACACTGCTCACCA 2700 Db
|||||
2701 GAGGGAACATTTTCACTGAGAGGTTCATGGGCTTACAGAGTGAACCTCGTATTTCTTC 2760 Qy
|||||
2701 GAGGGAACATTTTCACTGAGAGGTTCATGGGCTTACAGAGTGAACCTCGTATTTCTTC 2760 Db

Qy	3	GGCGCGGGCGGACACAGGGCCGCGCGGCTCTCGTGGTCTGACCTTCTGCGCTGCTGTGCTCCGCGCG	62
Db	219	GGGGGACGCGCGCGCGCGCGCTGGGCTCTCGCGTTGACCTTCTGCGCTGTGGCCGCGCG	278
Qy	63	CGGGGAGCTGCCATTGCGGCCACGAGGAGCAACTGTCTACAGCTGAGCTGTGTGATGAGGGACCCCT	122
Db	279	CGGGGAGCTGTGTGTGCGCCACGAGGAGCGACTGTGGAGCTGAGCTGTGGATGGGGGCACT	338
Qy	123	GCAAGTGATCTCGGGCCCTGAGCAGGCTGTGTGTCTGGACTGCACTTTGGGGGCTACAGC	182
Db	339	GCAAGTGATCTCGGGCCACGAGCAGGCTGTGTGTCTGGACTGCACTTTGGGGGCTGTGTC	398
Qy	183	TGCTGGGCTCCGACAGGATGACATGGAGCAAGGATGGAGACACTGTACTAGAGCATGTA	242
Db	399	CGCTGGACCCCCACCAAGGCTGACTTGGAGCAAGGATGGGAGACACCTGTCTGGAGCACGA	458
Qy	243	GMACCTGCACCTGTACTCCCAATGGCTCCCTGTGGCTGTCTCAACCTAGAGCAAGAGA	302
Db	459	CCACTTACCTGTCTGCCAATGGTCTCTGTGTGCTGTCCAGCCTAGCAACCCATGG	518
Qy	303	CAGCGATGATGAGGAAGCTCTTAGATCTGGAAAGGTCACTGAGGGCAGCTATTCTGTCT	362
Db	519	CAG---TGACGAGTCACTCCCTGAGGCTGTGGGGTCAATTGAAGGCAACTATTTCGTGCT	575
Qy	363	GGCCCAAGCCCGCTAGGAGTGTGGCCAGCAGTTGTGTGTCTCAAGCTTGTCCCACT	422
Db	576	AGCCCAAGGCCCCCTCGGAGTGTCTGGCCAGCAGACTGTCTGTCTCAAGCTTGTCCCACT	635
Qy	423	CGAGACTTCTCTGTGCACCCCGAGTCCAGATTGTGGAGAGAAACGGGACAGCAACGCTT	482
Db	636	CGCAGACTTCTCTGTGCACCCGAGTCTCAGACGGTGGAGAGAAACGGGACAGCTGCTT	695
Qy	483	TGAATGCCACACCAAGGCGCTTCCAGCGCCCATCATTTACTTTGGGAAAGACACAGGTGAC	542
Db	696	TGAGTGCAACTTGAAGGCTGCAGCTCCCATCATTTACTTTGGGAAAGACACAGGTGAC	755
Qy	543	CGTGCTGAGGAGCCCCGGCTCACTCTTTCCAGTGGCTCTCTCAGATCTCTAGATGT	602
Db	756	ATTGCTTGAGGAGCGCTCGGCTCATCTGTGCTTCCCAACGGCGTCTCTTTCAGATCTCTGATGT	815
Qy	603	CCAGGACAGTGATGACGGCTCTACCGCTGGGTGGCCACCAATTCAGCCCGCCCAACGATT	662
Db	816	TCAGGAGAGTGAAGAGGCCCTTACCCTCGTGGCCACCAACTCAGCTCGCCAGCACATT	875
Qy	663	CAGCCAGAGGCGCTCGTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTACACAGGGGCA	722
Db	876	CAGCCAGAGGCGCTCACTAGTGTGGCCCAAGAGGGTCCCTGGCGTCAACAGGGGCA	935
Qy	723	GGATGTGGTCAATTGTGGCAGCCCGAGAGAACACACAGGTAGTGTCTGGACAGAAATGTAGT	782
Db	936	GGACGTGGTCAATTGTGGCAGCCCGAGAGAACACACAGGTGTGTCTGGCCAGAGTGTGT	995
Qy	783	GATGGAGTGGTGGCGCTTGTGTACCCCAACCCCTTTGTCTCTGGGTCCGACAGGATGG	842
Db	996	GATGGAAATGTGTGGCGCTCAGCTGACCCCAACCCCTTTGTCTCTGGGTCCGACAGAGCGG	1055
Qy	843	AAAGCCTATCTCCACGAGTGCATCTGTTCTGGGGCGGACCAACTACTACTATCGCCAGCGC	902
Db	1056	GAAGCCCATCTTCAAGATGTCATCTGTGGCGGCAACCAACTACTATTATGCAACGC	1115
Qy	903	GCAGCTCGGCACTCTGGAGTCTATGTCTGCGAGCCAAACAGCCCCCTCAGCGGTGACTT	962
Db	1116	GCAGCCTTGGCACTCTCGGCGTCTATGTCTGCGCGGCCAAACAGCCCCCGCAACGCGACTT	1175
Qy	963	CGCCACTGCGGCTGCTGAGCTCCAGAGTGTCTGTGCCCGCAGCCATCTCGAGGCAACCGA	1022
Db	1176	CGCCACTGCAAGCGCTGAGCTCCGTGTGTCTGGCGGCTCCCGGCCATCACTCAGGCGCCCGA	1235
Qy	1023	GGCGCTCTCGGGGACGGGGGACAGCACGCGCGCTTCTGTGTGCGGGCGCTCCGGAGGCC	1082
Db	1236	GGCGCTGTGCGGGAACGGGGGACAGCACGCGCGCTTCTGTGTGCGCGCGCTGCGGGGAGCC	1295
Qy	1083	ACGGCCCGCGCTGCACTGGCTGCAAGAGGGATCCCGCTTGCAGACCCCAATGGGCGGCTCAA	1142

[illegible]

QY	4338	CCTGGTTTGGGAATCTTTTGGCAACCTGCTCTCTGCCCCCTTCTTCAATTTGTTT	4397
Db	4565	CTGGTTTGGAAATCCCGTTTCTCTCACTCTTCCCTCTTCTGCTGCTGCTGCTC	4624
QY	4398	TGGTGATCTGTCT - CATGTCACTTGTCTCGAATATC - CTGGGGCCCTTCTCTT - TCCCA	4453
Db	4625	TAGTGTCTGTCTCCAGTCACTGCTCTCCGACACAGTGGCTTCTCTGCTCCA	4684
QY	4454	TGATGCCCTGATTTCTCTCACTGCTGTTTTCAAT - TCTGTCTGCCATGCTTG	4504
Db	4685	GATACTCTTTCTTCTCTCTCTCTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCT	4744
QY	4505	TCTTTATGCTGTGTTTCTCTGCTCCCTGAGTTCAACCTATGACCCCTTCTCTCAACAT	4564
Db	4745	AGACTATGTCACTTGTCTCTGCTGGGTTCAAACTCTGCATCTCTCTCTCTCTCTCT	4804
QY	4565	GACTACCTCATGCTGTCTCTCA - GACCATAGTGTGACCCCTGGGTCCCCACAG	4615
Db	4805	GACTACCTCATGCTGTCTCTCAAGGCCCGGCTCTCTCTCTCTCTCTCTCTCTCTCT	4864
QY	4616	CTCCCTGCCAACCGCTTCTCT - GGGC	4641
Db	4865	CTGCTCTCCATCT	4924
QY	4642	AGATGAGCCCACTCAAGTAGATCTGGAAGAGACCTTGTGCTGCTGCTGCTGCTCTCT	4701
Db	4925	AGCCACCCAGCCAGTGGCGGTGTGGAAGATCTTCTGCTTCCCTGCTCT - CTGTC	4982
QY	4702	CTTTGGTGTGAGATGAGAGGTTTCTATGGAAGATGATCCAGGCTGCAAGGGA	4761
Db	4983	CTTTGGGATGGATCCAAAGGTTCTCCATGATG - GATCCAAAGTCATAGAGGGAA	5038
QY	4762	ACCCCAAGAGGGGTAGGAGTGAACCAAGAGGCTGAAAGAAATGGCTGCCACCC -	4819
Db	5039	TGTTTGACAGGGAAGGGGCTGTGATCCAGAGGCTCAGAAATAAAGATGCCCTCCT	5098
QY	4820	- ATCTGCAAGAGATGGGTGTGCTTTTGAAGTGCAGTCTCTGCTGGAACACTGAAGG	4878
Db	5099	TCTATGCAAGGGGGCAAGTTTACTGGATGGAGATGATTTGGGCTCTCTTCCAGAGAAG	5158
QY	4879	GTGAGGAGGGGAGCTACTGGGCTGCATGGCTCAGTCTCCCTGACCCCTGAGGCCCTGA	4938
Db	5159	CTAAGGAAGAGAGGGAGTGAAGTTCAGGGAGGCCCTTCCACCCCTGTGAGGCTTGA	5218
QY	4939	ACCTGGCTTCAGAGTAGCAAGAGTTT - CCTCAAGATGCTGAAGGAGTCTT	4992
Db	5219	CTTGATCTGGATTTGGGATGACAGGAATCTCACCTCTGGGGTGTGCAAGAGGTTCT	5278
QY	4993	TGCATAGGAAGGGCGCTGGCTCAATTTTATTTTATCTTTTCTTTTACCTGAATCCCAA	5052
Db	5279	TGCACAGGAAGG - GGTAGCTCATTTTCTAGTTTGTCTTTTAAATGAATCTCTCAA	5336
QY	5053	ATCAT - CTTACACAAAGGGCCAGCTGACTGCTGATTTCTGCTGAGTCAAA	5102
Db	5337	GTCAATTTTCTGTTCACCTTCCGCAAGGCTTGAATTTCTGCTGCTGCTGCTGCTGCT	5396
QY	5103	GAGCATGCCATCTCTGCTTCTCAGTCACTGCTCCAGATGCTCAGATCCAGT	5162
Db	5397	AAACAATGTCAATTTTGGTTTTCACCTCAGCCCTCTCATAGGATGATGATGTTAG	5456
QY	5163	GGCATCTGTCTTGTCTGCACATCTTTCTATTTTCAATGGCTGGCAATCAAGTGTAA	5222
Db	5457	GGTCTTTACTCCCTA - ATGATGCTGATTTGGCAATCAAGATGTA	5501
QY	5223	CTCTGGCTTCTGGGCAAGTAGAATAACAGTCTATTTTCC - TTTATTTATTTTA	5280
Db	5502	CTCTGCCCTTCTGGGCAATTCGAAATAACAGTCCATTTTCTCTTTTCTTTTCTTTT	5561
QY	5281	TTTTATTTATTTATGCTTCTCAGTGGAGTGTAGCTT - CTGAAGGCTCT	5331
Db	5562	TTTTAAATGGTGAATGCTCTCTCAGCACAGTGTGGGCTTCTCTCAAAACCCTGAAAGCATCT	5621
QY	5332	GTGTTTATTAGCTTGTGTGCTCACTCATGTTTGA - CCCCAAGCCACATTTCTCTCTCT	5388
Db	5522	GTGTTTATATATCTGCGGTGTCACTCACTGTGTGATGTCTGCACCTACGTTTCCACCTCT	5681
QY	5389	CCCC - TCTTCAGCAGCCTATGATAACACTAAAGATATTAATGCTGCTTCTGATCTC	5446
Db	5682	CCCCCTCTTCAGCAGCCTATGATAACACTAAAGATATTAATGCTGCTTCTGATCTC	5741
QY	5447	ATTAAGACAGGATGTCTCACTTGAATCTCTTATAGATTTCAAGTGGCCACGCCCAAC	5506
Db	5742	GTTAAGACAGGAATTTGCTCACTTGTGTATTTCTGTAGCATTTCAAGCTGCTGTGCTAAC	5801
QY	5507	ACCAC - CGTATGTTTCTCTCACTTGTCTGAGGCTCAAGAGCTCATTTTGTGCTGCTGT	5564
Db	5802	ACCATGTGTATGTTTCTCATCTGCTCTGAGGCTCAAGAGCTCATTTTATTTGCTGT	5861
QY	5565	TAGATTTCTTTCTCT - CCTTGCCTTGAATGAATAACCGTTTAAAC	5608
Db	5862	TTGATTTTCTTTTAAAGAGAAAGAAAGAAAGCTGCTGATTTAAATGGCTGTTTAAAC	5921
QY	5609	AGTAGGCTCTTAGCATCACACCATAGTCACTCTCATGTTCTTGTGTTTAAACAGCACTT	5668
Db	5922	AGTAGGCTCTTAGCATTATACCATAGTCACTTCTCATGTTCTTGTGTTTAAACAGCACT	5980
QY	5669	GAGGCTCTGGGTTTAAATTAATAGCTGCAATGAGACAAATTTATACCCATTAGGCTGG	5728
Db	5981	- GAGGCTCTGGTTTAAATTAATAGCTGCAATGAGACAAATTTATACCCATTAGGCTGG	6039
QY	5729	GTGGAATAATG - TTCTCAAGAGCAATAAGTAATAAATCTGCTATCTGCTTATACCTAC	5787
Db	6040	GTGGAATAATGTTTCTCAAGAGCAATAAGTAATAAATCTGCTATCTGCTTATACCTAC	6099
QY	5788	AGTTGATAAGAAAGTAGCCAGAACTCACTAGCATTATATATGATTTGGGTTCTGAGTAAC	5847
Db	6100	AGTTGATAAGAAAGTAGCCAACTTCTCACTAGCATTATATATGATTTGGGCTCTGGTAAT	6159
QY	5848	TGGGAGTGTAGCTTGTGATCTTGTGATGACACAGGCTCT - TATTAGGAAGATCTGTTGGC	5906
Db	6160	TTGGAAGTGTAGGTTTGTGTTTGTGATGAGTATTTTATAGGAAGATCTATTTGGC	6219
QY	5907	CTTTTACAGGGCATTTAGTCCCTTGTGTTTGGCATGAGTCCCTTAAAGTCTTTG - GAGT	5965
Db	6220	CTTTTACAGGGATTAATATCTTGTGATGATGATGATGATGATGATGATGATGATGATG	6279
QY	5966	CTCATTTAAGAAATCTCTTCTCGAAGCATGACAAAGTGTATGCAATCTTACATGCTCA	6025
Db	6280	CAATTAATAATCTCTTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATG	6334
QY	6026	CTCCTTTACCTG - GCTTAGTTTGTGCTGGGTTATTTAATGCACTTTCCAGCATGCT	6084
Db	6335	TTTATTTGTCTGATTTTGTGTTTGTGATGATGATGATGATGATGATGATGATGATGATG	6394
QY	6085	TCTCTCTTACAAATGATGATTTCTTTATTTGTTACACTAAGGTTTGA - TCAATGATCTGT	6143
Db	6395	TTCTCTTACAAATGATGATTTCTTTATTTGTTAAGCTAAGGATTTGATTTATGATCTGT	6454
QY	6144	CCCTGTAAGAAATTAATAAACTATTTTCCAG	6174
Db	6455	CTTATTAATGAATTAATAAACTATTTTCCAG	6485
RESULT 8			
AX191261			
LOCUS	AX191261	Sequence 3 from Patent WO0149714.	2796 bp DNA linear PAT 15-AUG-2001
DEFINITION	AX191261	Sequence 3 from Patent WO0149714.	
ACCESSION	AX191261		
VERSION	AX191261.1	GI:15209512	
KEYWORDS			
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
	Salbaum, J.M.		

TITLE	Nope polypeptides, encoding nucleic acids and methods of use				
	Patent: WO 0149714-A 3 12-JUL-2001;				
	Neurosciences Research Foundation Inc. (US)				
FEATURES	Location/Qualifiers				
	1..2796				
	/organism="Mus musculus"				
SOURCE	/mol_type="genomic DNA"				
	/db_xref="taxon:10090"				
	<1..>2796				
CDS	/notes="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAC51121.1"				
	/translation="GSLPQETTVKLSCEGPLOVILGPQAVLVLDCTLGATAAPP				
	TVWNSKQDGVLEHNLHLLPNSLWLSPLEQSDDEBALRIWKTEGYSCLAH				
	SELVGVSQAVVVKLALSDFLSPESQIVRENGTARFCHTKGLPAPITWKRQVTR				
	VPEEPLRLTQVLLQDSDAGSYECVATNSARQFSQEAISLVALRGSLEAIR				
	GDQVIVAPENITVVGQNVMECVASADPTFPVSWRQDKPISTDIVILGRNLL				
	IASQPRHSGVTVCRANKPLTRDFATAAELRLAAPASQAPALSRTRASTAFVC				
	RASGEPRPALHLDGIPLRPNRNVKVGSGSLVITQIGLODGYQCVASNSAGTA				
	CAAPLVAVVRGLPSAPTRVATPLSSSLVAVMRPELHSEQLIGFSLHYQKRGV				
	DNVEYFVANNDETEIQVRLDLPNTIDYFVVAISQIGASRTSSPALVHTLDDVFAA				
	PQLTSSPNPSDIRVLAWLPLPSLSNGQVTKYKIEYGLKEDQVSTFVPGNETQLT				
	NSLQNPVVRVIRISAGTGAGYGVPSQWMOHRTKPGVHNSHPVFAELKVRAMESLV				
	VSGQPPHPPTQISGYKLYWEGVTEBEADRPGRGQDQAVDVGVRUKKVKVQVEL				
	TOLVGRPYEVKLVAFNKHEGVAVWKCKTEKAPTDLPIORGPEPLPAHVHARSNS				
	STSLWRKWKPPFTTKIYNYTRFGPWGLRNALVITYTSSGREDILIGLAPPTKYE				
	FAVQSHGMDGPFVGSVERSLTPDRPSPDRLSLPTSPTRVLRHWCPTPEPNGEI				
	VELLYLSNNHTOPEHQWTLTTEGNIIFSAEVLHLESDFRYFFKMGARTFVGGPFPR				
	LDQVITLQETFSDSLVDH"				
BASE COUNT		591 a	821 g 535 t		
ORIGIN					
Query Match 45.3%; Score 2796; DB 6; Length 2796;					
Best Local Similarity 100.0%; Pred. No. 0;					
Matches 2796; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	64	GGGGAGCTGCATTGCCAGGAGACAACTGTCAAGCTGAGCTGTCAGAGGACCCCTG	123		
DB	1	GGGGAGCTGCCATTGCCAGGAGACAACTGTCAAGCTGAGCTGTCAGAGGACCCCTG	60		
QY	124	CAAGTGATCTGGGCGCTGAGCAGGCTGTGGTCTGGAGCTGCACTTTGGGGGCTACAGCT	183		
DB	61	CAAGTGATCTGGGCGCTGAGCAGGCTGTGGTCTGGAGCTGCACTTTGGGGGCTACAGCT	120		
QY	184	GCTGGGCTCCGACGAGGTGACATGGAGCAAGATGGAGACACTGTACTAGAGATGAG	243		
DB	121	GCTGGGCTCCGACGAGGTGACATGGAGCAAGATGGAGACACTGTACTAGAGATGAG	180		
QY	244	AACTGCACCTGTACCAATGCTCCCTGTGGCTGTCTCACCCTAGAGCAAGAGAC	303		
DB	181	AACTGCACCTGTACCAATGCTCCCTGTGGCTGTCTCACCCTAGAGCAAGAGAC	240		
QY	304	AGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTAGAGGCGAGCTATTCTGTCTG	363		
DB	241	AGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTAGAGGCGAGCTATTCTGTCTG	300		
QY	364	GCCACAGCCGCTAGAGGTGGTGGCAGCCAGGTTGCTGTGTCAAGCTTGCACACTC	423		
DB	301	GCCACAGCCGCTAGAGGTGGTGGCAGCCAGGTTGCTGTGTCAAGCTTGCACACTC	360		
QY	424	GAAGACTTCTCTGACACCGAGTCCAGATTGTGGAGGAGAACGGGACAGACGCTTT	483		
DB	361	GAAGACTTCTCTGACACCGAGTCCAGATTGTGGAGGAGAACGGGACAGACGCTTT	420		
QY	484	GAATGCCACACCAAGGCTTCCAGCCGCCCATCACTTACTTTGGGAAAGGACAGGTTGACC	543		
DB	421	GAATGCCACACCAAGGCTTCCAGCCGCCCATCACTTACTTTGGGAAAGGACAGGTTGACC	480		
QY	544	GTGCTGAGAGGCCCCGGCTCATCACTTTTCCAAAGTGGCTCTTCCAGATCTTAGATGTC	603		
DB	481	GTGCTGAGAGGCCCCGGCTCATCACTTTTCCAAAGTGGCTCTTCCAGATCTTAGATGTC	540		

QY	604	CAGGACAGTGATGACAGGCTCCTACCGCTGCGTGGCCACCAATTTCAGCCCGCCCAACGATTC	663
DB	541	CAGGACAGTGATGACAGGCTCCTACCGCTGCGTGGCCACCAATTTCAGCCCGCCCAACGATTC	600
QY	664	AGCCAGAGAGGCTCTGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTTACAGGGGGGAG	723
DB	601	AGCCAGAGAGGCTCTGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTTACAGGGGGGAG	660
QY	724	GATGTGGTCATTGTGGGAGCCCCAGAGAACCAACGCTAGTGTCTGGACAGAACTAGTGTG	783
DB	661	GATGTGGTCATTGTGGGAGCCCCAGAGAACCAACGCTAGTGTCTGGACAGAACTAGTGTG	720
QY	784	ATGGAGTGGTGGCTCTGCTGACCCACCCCTTTTGTGTCTCTGGGTCCGACAGGATGGA	843
DB	721	ATGGAGTGGTGGCTCTGCTGACCCACCCCTTTTGTGTCTCTGGGTCCGACAGGATGGA	780
QY	844	AAGCCTATCTCAACGATGTCACTGCTTGGGCGCGGACCAATCTACTATCATGCTCCAGCGG	903
DB	781	AAGCCTATCTCAACGATGTCACTGCTTGGGCGCGGACCAATCTACTATCATGCTCCAGCGG	840
QY	904	CAGCCTCGGCACTCTGGAGTCTATGCTCTGCGGAGCAACAAAGCCCTCAAGCGTCACTTC	963
DB	841	CAGCCTCGGCACTCTGGAGTCTATGCTCTGCGGAGCAACAAAGCCCTCAAGCGTCACTTC	900
QY	964	GCACCTGGGCTGTGAGCTCCGAGTGTCTGCTGCCCGCAGCCATCTCGCAGGACACCCGAG	1023
DB	901	GCACCTGGGCTGTGAGCTCCGAGTGTCTGCTGCCCGCAGCCATCTCGCAGGACACCCGAG	960
QY	1024	GGCCTCTCGCGAGCGCGGCGCAGACCGCGCGCTTGTGTGCGGGGCTCGCGGGAGGCA	1083
DB	961	GGCCTCTCGCGAGCGCGGCGCAGACCGCGCGCTTGTGTGCGGGGCTCGCGGGAGGCA	1020
QY	1084	CGGCGCGGCTGCACCTGGCTGCACACGCGGATCCCGTTGCGACCAATGGGGCGCTCAAG	1143
DB	1021	CGGCGCGGCTGCACCTGGCTGCACACGCGGATCCCGTTGCGACCAATGGGGCGCTCAAG	1080
QY	1144	GTGCAAGGCGGTGGGCGGAGCTTGTGCTCATCTCAGATCGGCTCTGCGAGACCTCGCTAC	1203
DB	1081	GTGCAAGGCGGTGGGCGGAGCTTGTGCTCATCTCAGATCGGCTCTGCGAGACCTCGCTAC	1140
QY	1204	TAACAGTGTGATAGCAAGAAAACAGCGCGGAACTGCTGTGCGGCTCGCGCCCTCGCGGTA	1263
DB	1141	TAACAGTGTGATAGCAAGAAAACAGCGCGGAACTGCTGTGCGGCTCGCGCCCTCGCGGTA	1200
QY	1264	GTGGTGGCGAGGGGCTGCCAGCGCCGCTCGAGCTCGGCTCAGACCGCGCTGAGCAGC	1323
DB	1201	GTGGTGGCGAGGGGCTGCCAGCGCCGCTCGAGCTCGGCTCAGACCGCGCTGAGCAGC	1260
QY	1324	TCCTCTGTGCTGGTGGCGGCTGAGTTGCAACGCGAGCAAAATCATTTGGCTTC	1383
DB	1261	TCCTCTGTGCTGGTGGCGGCTGAGTTGCAACGCGAGCAAAATCATTTGGCTTC	1320
QY	1384	TCCTCTCACTCAAAAGGCAAGGGAGTGGACAATGTGGAGTACACGATTTGCACTAAAC	1443
DB	1321	TCCTCTCACTCAAAAGGCAAGGGAGTGGACAATGTGGAGTACACGATTTGCACTAAAC	1380
QY	1444	AAATGACACACAGAGCTCGAGGTTGGGACCTGGACCCCAACAGGATTTAGTTCTAC	1503
DB	1381	AAATGACACACAGAGCTCGAGGTTGGGACCTGGACCCCAACAGGATTTAGTTCTAC	1440
QY	1504	GTGGTGGCTTACTCCAGCTGGGGCGCAGCCGAACTCCAGCCAGCCCTGGTGCATACA	1563
DB	1441	GTGGTGGCTTACTCCAGCTGGGGCGCAGCCGAACTCCAGCCAGCCCTGGTGCATACA	1500
QY	1564	CTGGACGATGTCCCAGCGCAGACCCAGCTTACTTGTTCAGGCCCAACCCCTCGGAC	1623
DB	1501	CTGGACGATGTCCCAGCGCAGACCCAGCTTACTTGTTCAGGCCCAACCCCTCGGAC	1560
QY	1624	ATCAGGGTGGCATGGCTGCCCCCTCCAGCTGAGCAATGGACAGGTGCTGAAGTAC	1683
DB	1561	ATCAGGGTGGCATGGCTGCCCCCTCCAGCTGAGCAATGGACAGGTGCTGAAGTAC	1620
QY	1684	AAGATAGATGACGTTTGGGGAGGAAGATCAGGTTTTTCTCCACCGAGGTGCTTGGAAAT	1743

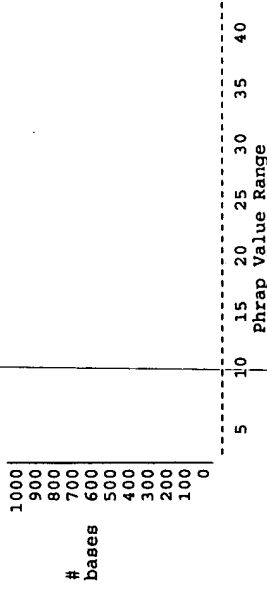
Genes and Regions of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to EST and cDNA sequences in Unigene. Genes demonstrate at least two exons flanked by consensus splice sites that maintain sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Attempts are made to complete double stranded sequence for all regions. All sequence is completed to a standard of coverage with a minimum of 3 reads with no ambiguities. If the sequence coverage for a region does not meet this standard, it is indicated in the annotation as Low Coverage. Low coverage linkages are verified by PCR product size verification or verification of forward and reverse reads from clones which span the low coverage area.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated average error rate is less than 1 per 10,000 bases using the Consed quality parameters. Regions that do not meet this requirement are annotated as Low Quality.

-----Summary Statistics
Center project name: ALH
Sequencing vector: pSMART; AF399742
Chemistry: Dye-terminator Big Dye; 100%
Assembly program: Phrap version 0.990319
Contig length: 115035
Fraction of Phrap value < 40: 0
Error Rate in Consed: 0.00 per 10,000 bases
Number of N's in consensus: 0

----- Distribution of Quality < 40 Bases:



FEATURES	Source	Location/Qualifiers									
		1	5	10	15	20	25	30	35	40	
		Phrap Value Range									

		Location/Qualifiers									
		1. .115135									
		/organism="Mus musculus"									
		/mol_type="genomic DNA"									
		/strain="ES cell line"									
		/db_xref="taxon:10090"									
		/chromosome="9"									
		/clone="MGS1-117K9"									
	repeat_region	190..274									
		/rpt_family="ID_B1"									
	repeat_region	275..497									
		/rpt_family="B3"									
	repeat_region	498..652									
		/rpt_family="ID_B1"									
	repeat_region	760..829									
		/rpt_family="L2"									
	repeat_region	870..990									
		/rpt_family="ORR1D"									
	STS	1173..1422									
		/standard_name="C79952"									
	repeat_region	2880..2914									
		/rpt_family="(TG)n"									
	repeat_region	complement(2918..3013)									
		/rpt_family="RSINE1"									
	repeat_region	3119..3173									

Query Match 44.4%; Score 2743.2; DB 10; Length 115135;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2818; Conservative 0; Mismatches 13; Indels 7; Gaps 6;
QY 3341 AGAGCAACGGAGAGAGAGCGCTCCCGAGGAGTCAAGTGGAGCTGAGGTCA 3400
Db 105112 AGAGCAACGGAGAGAGAGCGCTCCCGAGGAGTCAAGTGGAGCTGAGGTCA 105053
QY 3401 TTGTCCACTCCGACTTCGGTGCATCCAAAGGATGCTCCTGACCTCCACCTCCAGACCTGG 3460


```
QY 5621 GCATCACACACATAGTCACTTCCTCATGTTCTTGTTTAAACAAGCACTTCGAGGTTCTGGGT 5680
|||
Db 102834 GCATCACACACATAGTCACTTCCTCATGTTCTTGTTTAAACAAGCACTTCGAGGTTCTGGGT 102777
|||
QY 5681 TTAATAATTAATAGTCTCAATAGTACAGCAATTTATTAACCCATTAAGGCTGGGTGGAAATTTG- 5739
|||
Db 102776 TTAATAATTAATAGTCTCAATAGTACAGCAATTTATTAACCCATTAAGGCTGGGTGGAAATTTGT 102717
|||
QY 5740 TTCTCAAAAGCAATTAAGTAAATTAATCTGGTATCTGCCTATACTCACAGTTGATAGA 5799
|||
Db 102716 TTCTCAAAAGCAATTAAGTAAATTAATCTGGTATCTGCCTATACTCACAGTTGATAGA 102657
|||
QY 5800 AGTAGCCAGCACTCCTAGCATTATATATGATGCGGTTCTGAGTAACCTGGGGAGTGTTA 5859
|||
Db 102656 AGTAGCCAGCACTCCTAGCATTATATATGATGCGGTTCTGAGTAACCTGGGGAGTGTTA 102597
|||
QY 5860 GCTTTGTGACTTTGTAGCACCAGTCTTATTAAGGAAAGTCTGTTGCCCTTTTACAGGGCA 5919
|||
Db 102596 GCTTTGTGACTTTGTAGCACCAGTCTTATTAAGGAAAGTCTGTTGCCCTTTTACAGGGCA 102537
|||
QY 5920 TTAGTCCCTTTGCTGTTGGCATGATGCGCTTAAGTCTTTGAGTCTCATTTAAGAAAT 5979
|||
Db 102536 TTAGTCCCTTTGCTGTTGGCATGATGCGCTTAAGTCTTTGAGTCTCATTTAAGAAAT 102477
|||
QY 5980 CTTTTTCTCGAAGCATGACAAAGTGTATCGCACTTACATGCTCACTGTTTACCTG-G 6038
|||
Db 102476 CTTTTTCTCGAAGCATGACAAAGTGTAT-GCAATCTTACATGCTCACTGTTTACCTG 102418
|||
QY 6039 CTTAGTTTGTGCTGGGTATTTAATTTGCACTTTCCAGCATCATGCTTCTCTCTTACAAAT 6098
|||
Db 102417 CTTAGTTTGTGCTGGGTATTTAATTTGCACTTTCCAGCATCATGCTTCTCTCTTACAAAT 102358
|||
QY 6099 ATGATATCTTTATTTGTACACTAAGTGTGATCATGATCTGTCCTCTTAAGAAATTA 6158
|||
Db 102357 ATGATATCTTTATTTGTACACTAAGTGTGATCATGATCTGTCCTCTTAAGAAATTA 102298
|||
QY 6159 ATAACTATTTTCCAGAC 6176
|||
Db 102297 ATAACTATTTTCCAGAC 102280
|||
```

RESULT 10

```
AC110235/c
LOCUS AC110235 158055 bp DNA linear HTG 15-MAY-2002
DEFINITION Mus musculus clone RP23-100M12, WORKING DRAFT SEQUENCE, 17 ordered
pieces.
ACCESSION AC110235.5 GI:20800205
VERSION HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 158055)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE Birren, B., Linton, L., Nusbaum, C. and Lander, E.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 158055)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliiev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C.,
Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T.,
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C.H., O'Connor, T., O'Donnell, P., and Lander, E.
```

```
Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
Rovetti, P., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S.,
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (10-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 158055)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L.,
Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliiev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L.,
Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (15-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 15, 2002 this sequence version replaced gi:20531900.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L22383
Center clone name: 100_M.12
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 153807 bases at least Q40
Consensus quality: 155134 bases at least Q30
Consensus quality: 155784 bases at least Q20
Insert size: 218000; agarose-fp
Insert size: 156455; sum-of-contigs
Quality coverage: 7.1 in Q20 bases; agarose-fp
Quality coverage: 10.0 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 6459: contig of 6459 bp in length
* 6460 6559: gap of 100 bp
* 6560 6667: contig of 108 bp in length
```

```

* 6668 6767: gap of 100 bp
* 6768 7863: contig of 1096 bp in length
* 7864 7963: gap of 100 bp
* 7964 9426: contig of 1463 bp in length
* 9427 9526: gap of 100 bp
* 9527 11011: contig of 1485 bp in length
* 11012 11112: gap of 100 bp
* 11113 12741: contig of 1630 bp in length
* 12742 12841: gap of 100 bp
* 12842 14789: contig of 1948 bp in length
* 14790 14889: gap of 100 bp
* 14890 17826: contig of 2937 bp in length
* 17827 17927: gap of 100 bp
* 17928 21244: contig of 3318 bp in length
* 21245 21345: gap of 100 bp
* 21346 27413: contig of 6069 bp in length
* 27414 27513: gap of 100 bp
* 27514 36381: contig of 8868 bp in length
* 36382 36481: gap of 100 bp
* 36482 45104: contig of 8623 bp in length
* 45105 45204: gap of 100 bp
* 45205 58369: contig of 13165 bp in length
* 58370 58469: gap of 100 bp
* 58470 85782: contig of 27313 bp in length
* 85783 85882: gap of 100 bp
* 85883 112706: contig of 26824 bp in length
* 112707 112806: gap of 100 bp
* 112807 153927: contig of 41121 bp in length
* 153928 154027: gap of 100 bp
* 154028 158055: contig of 4028 bp in length.
FEATURES
    source
        1..158055
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /clone="RP23-100Mi2"
            /clone_lib="RPCI-23 Female Mouse BAC"
        1..6459
            /note="assembly_fragment"
            /clone_end:SP6
            /vector_side:left"
        6560..6667
            /note="assembly_fragment"
        6768..7863
            /note="assembly_fragment"
        7964..9426
            /note="assembly_fragment"
        9527..11011
            /note="assembly_fragment"
        11112..12741
            /note="assembly_fragment"
        12842..14789
            /note="assembly_fragment"
        14890..17826
            /note="assembly_fragment"
        17927..21244
            /note="assembly_fragment"
        21345..27413
            /note="assembly_fragment"
        27514..36381
            /note="assembly_fragment"
        36482..45104
            /note="assembly_fragment"
        45205..58369
            /note="assembly_fragment"
        58470..85782
            /note="assembly_fragment"
        85883..112706
            /note="assembly_fragment"
        112807..153927
            /note="assembly_fragment"
        154028..158055
            /note="assembly_fragment"

```

```

clone_end:T7
vector_side:right"
BASE COUNT 41469 a 36830 c 36521 g 41635 t 1600 others
ORIGIN
Query Match 44.4%; Score 2743.2; DB 2; Length 158055;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2818; Conservative 0; Mismatches 13; Indels 7; Gaps 6;
QY 3341 AGAGCAACGGGAGAGAAAGCCGTCGCCAGCGCTGCAGGAATCAGGTGGAAGCTGAGGTCA 3400
DB 73684 AGAGCAACGGGAGAGAAAGCCGTCGCCAGCGCTGCAGGAATCAGGTGGAAGCTGAGGTCA 73625
QY 3401 TTGTCCACTCCGACTTCGGTGCATCCAAAGATGTCCTGACCTCCACCTCCAAAGACCTGG 3460
DB 73624 TTGTCCACTCCGACTTCGGTGCATCCAAAGATGTCCTGACCTCCACCTCCAAAGACCTGG 73565
QY 3461 AGCCAGAGGACCACTGACTGCGAGAGACTCTGCCTTCCAGCTCTGGAGCTGTGATCTGT 3520
DB 73564 AGCCAGAGGACCACTGACTGCGAGAGACTCTGCCTTCCAGCTCTGGAGCTGTGATCTGT 73505
QY 3521 CTCAGGAGCAGACTGGCTGGGCGAGGAGCTGGGAGGGTGCACCAACCAACCAAGCTGGC 3580
DB 73504 CTCAGGAGCAGACTGGCTGGGCGAGGAGCTGGGAGGGTGCACCAACCAACCAAGCTGGC 73445
QY 3581 CAGAGAGCTCACCTGCTTGCAGAGCAGCCAGTGCCTCTGCTCTGCTCAGACCTCC 3640
DB 73444 CAGAGAGCTCACCTGCTTGCAGAGCAGCCAGTGCCTCTGCTCTGCTCAGACCTCC 73385
QY 3641 AGCCAGCAGCTCTATAGAGAGGCGCTTGGGAAAGCTGCGAGCCCAAGCCCTGTGTC 3700
DB 73384 AGCCAGCAGCTCTATAGAGAGGCGCTTGGGAAAGCTGCGAGCCCAAGCCCTGTGTC 73325
QY 3701 CTCTAAACAGTCAGCCCAAGCCTTCCAGGGCCCTCTGCTCTCTCTGCTCAGGTCCCTCAG 3760
DB 73324 CTCTAAACAGTCAGCCCAAGCCTTCCAGGGCCCTCTGCTCTCTCTGCTCAGGTCCCTCAG 73255
QY 3761 CAGAAGGCAGATATGGCTCAGGAACATGCCATGCAATGCTTACATATGTTGTTACTAGAGA 3820
DB 73254 CAGAAGGCAGATATGGCTCAGGAACATGCCATGCAATGCTTACATATGTTGTTACTAGAGA 73205
QY 3821 TATCCATAAGTCTTGGAGCCTCTTAGGGTCTTTGGCTGGGGTGGGGAGAACCTTTACT 3880
DB 73204 TATCCATAAGTCTTGGAGCCTCTTAGGGTCTTTGGCTGGGGTGGGGAGAACCTTTACT 73145
QY 3881 CTCCCTCATATTTGTCATCATACAGGAGGACCTTGAGACACACAGCTCTGTGTATGATGAC 3940
DB 73144 CTCCCTCATATTTGTCATCATACAGGAGGACCTTGAGACACACAGCTCTGTGTATGATGAC 73085
QY 3941 ACGTGTGAAGTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4000
DB 73084 ACGTGTGAAGTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 73025
QY 4001 CCTATGTAGCACTCACTGTGGCTTAGTTGACCCCTCGGTGGCAGGATGTTGTAACAGTGTAT 4060
DB 73024 CCTATGTAGCACTCACTGTGGCTTAGTTGACCCCTCGGTGGCAGGATGTTGTAACAGTGTAT 72965
QY 4061 CAGTGCCAGCTCTTTGAGCTTTTACCTTGTGCACTAGCCTTTATTTACATCTGAGAGT 4120
DB 72964 CAGTGCCAGCTCTTTGAGCTTTTACCTTGTGCACTAGCCTTTATTTACATCTGAGAGT 72905
QY 4121 GTCTCCAGTGTGTGTCTACAAAGACAGCGCCGACCCCTCTTCTGTCTGAGCTGTGCTGAGC 4180
DB 72904 GTCTCCAGTGTGTGTCTACAAAGACAGCGCCGACCCCTCTTCTGTCTGAGCTGTGCTGAGC 72845
QY 4181 AGAGTGCCAGTCAACTCCACCGGCTTATGACACCGAGCCTTACCAACAGCATGGCTGTGCAT 4240
DB 72844 AGAGTGCCAGTCAACTCCACCGGCTTATGACACCGAGCCTTACCAACAGCATGGCTGTGCAT 72785
QY 4241 CCCCCTGGCCTCCTAAGGTCAGATGTTCTGGGTGAAACCCAGCTCAGCTCCCTCTCCCTTT 4300
DB 72784 CCCCCTGGCCTCCTAAGGTCAGATGTTCTGGGTGAAACCCAGCTCAGCTCCCTCTCCCTTT 72725

```


QY 2371 CTCAGGAATGCTTCCCTGGTCACTACTATACAGCTCTGGAGAGACATTTCTCATTTGGC 2430
Db |||||
QY 1644 CTCAGGAATGCTTCCCTGGTCACTACTATACAGCTCTGGAGAGACATTTCTCATTTGGC 1703
Db |||||
QY 2431 GGCCTGAACCAATTTACCAAGTACGATTTGGGTGAGTCCACGAGATGATGAT 2490
Db |||||
QY 1704 GGCCTGAAGCCCAATTTACCAAAATACGAGTTTGCAGTGTGAGTGTGAGTGTGAT 1763
QY 2491 GGCCTTTGGCTTCCGCTGTGAGAACCTCCACCTGCTGACCGGCTTTCAACACTCT 2550
Db |||||
QY 1764 GGCCTTTTCCGCTTCTGCTGTGAGAGCTCCACCTGCTGACCGGCTTCCACACCCCA 1823
QY 2551 TCTGACCTGCGCTGAGCCCTCTGACACCATCCACCGTTCCGTTTCACTGCTGCTCCCT 2610
Db |||||
QY 1824 TCCGACCTGGACTGAGCCCTCTGACACCGTCCACCGTTGCGTGTGCTGCTGCTCCCT 1883
QY 2611 ACGAGCCCAATGCTGAGATTTGAGATTTCTAAATTTCTTACAGCAACCAACACCCAG 2670
Db |||||
QY 1884 ACAGAGCCCAACCGGGAGATCGTGGAGTATCTGATCTGTACAGCAGCAACCAACGCGAG 1943
QY 2671 CCGACACACAGTGGACATCTGCTCACACAGAGGGAACATCTTCAAGTGCAGAGTCCAT 2730
Db |||||
QY 1944 CTTGAGCACAGTGGACCTTGTCTCACACGACGGAACATCTTCAAGTGTGCTGAGGTCCAT 2003
QY 2731 GGCCTAGAGATGACACTCGGTATTTCTTCAAGATGGAGCCCGCACAGAGGTGGGCT 2790
Db |||||
QY 2004 GGCCTGAGAGGACACTCGGTACTTCTTCAAGATGGGGCGGCACAGAGTGGACCT 2063
QY 2791 GGCCTTTTCCGCTTGCAGGATGATTTACTTGTCAAGAGACATTTCTAGACTCTCTTG 2850
Db |||||
QY 2064 GGCCTTTTCTCCGCTGCGAGATGTATCAAGCTCCAGAGAACTGTGAGACTGCTGCTG 2123
QY 2851 GATGTGACCGCGTCAAGGACATCTGCTGGGTGTCTGCTGGCTTCTCTGCTCTCTG 2910
Db |||||
QY 2124 GACATGCACTAGTCAAGGACATCTGCTGGGTGTCTGCTGGCTTCTCTGCTCTCTG 2183
QY 2911 GCTGCACTGTGCTGGCTTACGACAAAGCTTCCACAGGAGAGCCCTTCCCGATGTGCT 2970
Db |||||
QY 2184 GCTGCACTGTGCTGGCTTGTGCGCGAGCCCGCACAGGGAATCTTCCAGGCTGTGCT 2243
QY 2971 TCTCAGGCACTCCAGGAAACCGAGCTTCTACAAAGAGCTGGGTGGGCTTCCAGT 3030
Db |||||
QY 2244 TCCACCGCACCCCGGAACTCCCGGCTGTACTCCAGAGCTCGGCTTGGCCCGCCAGC 2303
QY 3031 GTCCCTGTGCTGATGAGTGGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3090
Db |||||
QY 2304 CCCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2363
QY 3091 CCACCTCAGATGTGGAAGACAGGCTGAAAGTACAGAGCTTATGGGTGGCAGTGTTC 3150
Db |||||
QY 2364 CCACCTCAGAGTGTGGAAGACAGGCTGAAAGTACAGAGCTTATGGGTGGCAGTGTTC 2423
QY 3151 GATTCGGGGCACTCCAAGAGAAAGATCTCTGGGCTCAGGAGGGGGAACCAACTGG 3210
Db |||||
QY 2424 GAAGCGCGAGTCACTCCAAGAGAAAGATCTCTGGGCTCAACCAAGCGGCTGAGCTG 2483
QY 3211 GCAGGCTCTGGGAGGCTGTGAGCTGCCCCAGGATGTGCTGCTGCTGCTGCTGCTG 3270
Db |||||
QY 2484 GCTGGTTCCTGGGAGGCTGTGAGCTGCCCCAGG---CAGGGCCCCGGGCTGAGC 2540
QY 3271 CGTGTCTGCTGCTCAGCGGGAACCGGCGACAGACTGCTGCTGCAAGCCCTGGGTAT 3330
Db |||||
QY 2541 CGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2600
QY 3331 GACGGCAATGAAGCAACGGGAGAAAGAGCGCTCCCGAGCTGCGAGGAATCAGGTGGAA 3390
Db |||||
QY 2601 GACGGCAATGAAGGCAATGGGAGAAAGAGTCAACCCCGAGCTGCGAGGAACAGGTGGAG 2660
QY 3391 GCTGAGGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3450
Db |||||
QY 2661 GCTGAAGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2720

QY 3451 CAAGACCTGGAGCCAGAGGAACCACTGACTGACAGAGACTCTGCTTCCACCTCTGAGACT 3510
Db |||||
QY 2721 CAAGACCTGGAGCCCTGAGGACCCCTGCTCCTCAGAGGCTCTGATCTCATCTCGGTGT 2780
QY 3511 GTGATCTGTCTCAAGGAGCAGACTGCTGCTGGGAGGAGCTGGGAGGAGTGCACCAACA 3570
Db |||||
QY 2781 GGGGATCCAGGGCAGGGGAGCTGCTGCTGACAGGAGTGTGGAGGAGTGTGAGCTGCA 2840
QY 3571 ACCAGTGGGCGAGAGAGCTCACTGCTTGGCAGAGCAGCAGCTGCTCTCTGCTCTGCT 3630
Db |||||
QY 2841 GCGGCGGCGCAGACAGACTTACCTGCTTGGCAGAGGAGCAGCTGCTCTCTGCTCTAC 2900
QY 3631 TCAGACCTTCCAGCCAGCAGCTGCTATAGAGAGGCGCTTGGGAAAGCTGCGAGCCAAA 3690
Db |||||
QY 2901 CCGGACCTTCCAGCCAGGCGAGTGTAGAGGAGACCCCTGAGATAGCTGCGCAGCTCAA 2960
QY 3691 GCGCTGTGCTCTTAAAGTCAAGTCAAGCCCTTCCAGGCGCTGCTGCTCTCTGCTCTGCT 3750
Db |||||
QY 2961 TCCGCTGCTCTTAGGAGCCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3020
QY 3751 GTCCCTGAGCAGA-----AGGCGATATGGCTCAGGAACAT----- 3787
Db |||||
QY 3021 TAGCTTCTCCAGAGGATGTGCTTGGGCGAGGAGTATGATCACTAGGATGCGATA 3080
QY 3788 -----GCCATGCAATGCTACACATGTGTACTAGAGATATCCATAGTCTCTTGGAGCC 3841
Db |||||
QY 3081 CCGTGGCGGTGTATGTCCACATGTGTGCTGTAGATACATCATCAAGCCCTTTGGAGCT 3140
QY 3842 TCTTAGGCTCTTTGGCTGGGGTGGGAGAA-----CTTTACTCTCTCTCTATTTCTG 3895
Db |||||
QY 3141 TCTTAAGTGTCTTTGGCTGAGGGGAGAGAAACATGGATTTATCACTCCCCCCTATCTC 3200
QY 3896 CATCACATACAGGAGGACTTGGACACAGCTCTGTGTAAATGGACACCTGTGAAAGTCTGT 3955
Db |||||
QY 3201 TTTGTGATACATGTGACATGTGAAGACATACAGACATAGCTATCATGTATGTGAC 3260
QY 3956 TGT 4010
Db |||||
QY 3261 ATGTGTGAAGTGTATGT 3320
QY 4011 ACTCACTGTGGCTTGTGACCTTCCGCTGGGAGGATGTGTGTGTGTGTGTGTGTGTGT 4067
Db |||||
QY 3321 GGTCACTACGCTGT 3380
QY 4068 AGCTCTTTGAGCTTTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 4125
Db |||||
QY 3381 AGCCCTGTGGGCCCCCCTCCTGCTGCTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCT 3440
QY 4126 CAGTGTGTGTGTCTCAAAAGACAGCGCCCTCTTGTGTGCTGCTGCTGCTGCTGCTGCT 4185
Db |||||
QY 3441 CAATGCTGTGTGACAAAGACAGTCTCCAGCCCAATCTCTCTGCTGCTGCTGCTGCT 3497
QY 4186 GCGAGTCAACTCCAGGGGCTATGACACCGAGCTTACACAGATGTGTGTGTGTGTGTGT 4245
Db |||||
QY 3498 GCAA-----GCAAGGCTCTGAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3532
QY 4246 TGGGCTCTTAAGTCCAGATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4305
Db |||||
QY 3533 ATCACTCCAGCTCTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3592
QY 4306 TCTGTGTACCTAAATTTGTAAATCTGGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4365
Db |||||
QY 3593 ACTCTGCAATTTAAATTTGTAAATCTGGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 3652
QY 4366 CTGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4424
Db |||||
QY 3653 CTCTTCT 3712
QY 4425 CGATTTATC-CTGGGCGCTCTCTCTT--TCCCATGATGCTGCTGCTGCTGCTGCTGCT 4481
Db |||||
QY 3713 CTCCAGCAGCAGTGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3772
QY 4482 TTCAAT-----TCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4532

Db 3773 TTCTTCCTCTGCTATCTCCTCAGACCTCTCCAGACTATGTCATCTGTTCTCTCTGCTG 3832
Qy 4533 AGTTCAAGCTATGACACCTTTCCTTAACAAATGACTACCTCATGTCGCTTCA----- 4585
Db 3833 GGTTCNAATCTGCACTCTTCTTAACAACTGACTACCTCATGTCGCTTCAAGGCCCC 3892
Qy 4586 --GACCAATAGTGTGACCCCTGGGTCCCAAGCTCCCTGCGCAACCGCTTCCT----- 4637
Db 3893 CGTGCCTTCTGATCGCGGCTGCGCGCACTCGCGCTGCACTCTCTCTGCTCTCTCT 3952
Qy 4638 -----GGCAGATGAGCCCACTCCAGTAGATCTGGA 4669
Db 3953 CACTCAGTGTCTGCTTGCCTGCGCCCTGCGCCAGCGAGCCACCCAGCTGCGGGTGG 4012
Qy 4670 AAAGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4729
Db 4013 AGRAGATCTTCTGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4070
Qy 4730 ATGAAGAGATGATGCTCAGGCTGCAAGGGGAAACCCCAAGAGGGGTAGGGAGTGAAC 4789
Db 4071 ATGGATG---GATCCAAGTCAATAGAGGGGAATGTTTGAGACAGGGAGGGGGCTGTGAT 4126
Qy 4790 CAAGGCTGAAAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4846
Db 4127 CCAGAGGCTCAGAAATAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4186
Qy 4847 TTTGACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4906
Db 4187 GGAGATGTTGGGCTCTCTTCAGAGAGAGCTAAAGAGAGAGAGAGAGAGAGAGAGAGAT 4246
Qy 4907 CATGGCTCAGTTCCTGACCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTT- 4965
Db 4247 CAGGAGGCTTCCACCTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 4306
Qy 4966 -----CCTCAAGATGCTGTAAGGAGTCTTTGATAGGAGGAGGAGGAGGAGTCAAT 5020
Db 4307 CTCACCTCTGGGTGCTGCGAGGAGTCTTTGCAAGGAGGAGGAGGAGGAGGAGGAGG 4364
Qy 5021 TTATTTATCTTCTTACCTGATCCCAAGTCT-----CTTACCAAGG 5070
Db 4365 CAGTTGTTTCTTTAAATGAACTCTCAAGTCAATTTCTGTTCACTGCGCCAGG 4424
Qy 5071 GCCAAGCTGACTGCTATTTCTGAGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5130
Db 4425 GACAAGCTTGACTTCTATTTCTGCTAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4484
Qy 5131 CTCAGTCAATGCCAGAGTGTGAGTCCAGTGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5190
Db 4485 CTCAGCCTCTCATAGAGCATAGAAATGTTAGGGTCTTTACTCCCTA----- 4531
Qy 5191 CTATTTCAAGCTGCTGCAATCAAGTGTAACTCTGCTTCTGGCCAGGTTAGAATA 5250
Db 4532 --ATGATGCTGATGAGGAGATCAAGAGTAACTCTGCTTCTGGCCAAATTCGAAATA 4589
Qy 5251 ACCAGTCTATTTTCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 5304
Db 4590 ACCAGTCAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTCT 4649
Qy 5305 GTGAGTGTAGCTT-----CTGAAAGCTGCTGTTTATTTATTTATTTATTTATTTATTTATTT 5355
Db 4650 GCAGTGTGGCTTCTCCTCAACCTGAAAGCATCTGTTTATTTATTTATTTATTTATTTATTTATTT 4709
Qy 5356 TCATGTTGA---CCCAAGCAGATTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5410
Db 4710 TCATGTTGATGTGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4769
Qy 5411 TAACACTAAAGATTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5470
Db 4770 TAACACTAAAGATTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4829
Qy 5471 ACTACTTCTATGATTTCAAGTGGCCACCGCCCAACCACT--CGTATGTTTCTTCAATG 5528

Db 4830 AGTATTTCTGTAGCATTCAGCGCTGCTGTGGCTAAACACCACTGCTGTATGTTTTCATCATG 4889
Qy 5529 CTCTGAAGGTCAAGAGCTCATTTTGTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT 5581
Db 4890 CTCTGAAGGTCAAGAGCTCATTTTATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 4949
Qy 5582 -----TGCCTTGAATGAATAAACCCTTTTAAACAGTAGGCTCTTAGCATCACACAC 5632
Db 4950 AAAAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5009
Qy 5633 ATAGTCAATTTCTCATGTTCTTGTGTTTAAACAGCACTGAGGTTCTGGGTTTAAATTAATA 5692
Db 5010 ATAGTCAATTTCTCATGTTCTTGTGTTTAAACAGCACT--GAGGTTCTGCTGTTTAAATTAATA 5067
Qy 5693 GCTGCAAAATGAGCAAAATTTAATACCACTAGGCTGGGTGCAAAATTTG-TTCTCAAAAGCA 5751
Db 5068 GCTGCAAAATGAGCAAAATTTAATACCACTAGGTTGGGTGCAAAATTTGTTTCTCAAAAGCA 5127
Qy 5752 AATAAGTAATAATCTGCTATCTGCTATACCTACAGTTGATAAGAAAGTACGACCAAC 5811
Db 5128 AATAAGTAATAATCTGCTATCTGCTATACCTACAGTTGATAAGAAAGTACGACCAATTC 5187
Qy 5812 TCAGTACATATATATGATTGGGTTCTGAGTAACTGGGAGTGTAGCTTTGTGCTT 5871
Db 5188 TCAGTACATATATATGATTGGGTTCTGAGTAACTGGGAGTGTAGCTTTGTGCTT 5247
Qy 5872 TGTAGCACCAAGGCT-TATTAGGAAAGTCTGTTGGCTTTTACAGGCACTAGTCCCTTT 5930
Db 5248 TGTAGCAAGTATTTTATAGAAAGATCTATTGGCTTTTACAGGATATTAATCCCTTT 5307
Qy 5931 GTGCTTTGCGATGCTTAAAGTCTTTG-GAGTCTCATTTAAAGAAATTCCTTTCTCG 5989
Db 5308 GTCACTTACATTTGATGCTTAAAGTCTTCTGAGTCTCAATTAAGAAATTCCTTTCTG 5367
Qy 5990 AAGCATGCAAGTGTATCGCAATCTATCATGCTACTCTCTTACCTG-GCTTAGTTGT 6048
Db 5368 ATGATGACCAAGTGT-----AATCAGTACTTGTCTCATTTATTTCTGTTATTTAGTTAT 5422
Qy 6049 GCTGGGTATTTAATGCACTTTCCAGCATCATGCTTCTCTTACAAATATGATATCT 6108
Db 5423 GCTGTACTATTTAATATCTTCCAGGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5482
Qy 6109 TTATTTGTTACATAAGGTGTTGA-TCATGATCTGCTCTCTGTTAAAGAAATTAATAACTAT 6167
Db 5483 TTAGTGTAGCTAAGGCAATGATTCATGATCTGCTCTTATATGATTAATGAATTAATAACTAT 5542
Qy 6168 TTTCCAG 6174
Db 5543 TTTCCAG 5549

RESULT 12
AX418406
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .3741
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
/note="NOVX nucleic acid with homology to NOPE-like"

AX418406
Sequence 1 from Patent WO0206329.
AX418406
AX418406.1 GI:21523336
unidentified
unidentified
unclassified.
1
Rastelli, L., Shimkets, R.A., Zerhusen, B., Malyankar, U.M. and
Padigaru, M.
Human polynucleotides and polypeptides encoded thereby
Patent: WO 0206329-A 1 24-JAN-2002;
Curagen Corporation (US)
Location/Qualifiers
1. .3741
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
/note="NOVX nucleic acid with homology to NOPE-like"

AX418406
Sequence 1 from Patent WO0206329.
AX418406
AX418406.1 GI:21523336
unidentified
unidentified
unclassified.
1
Rastelli, L., Shimkets, R.A., Zerhusen, B., Malyankar, U.M. and
Padigaru, M.
Human polynucleotides and polypeptides encoded thereby
Patent: WO 0206329-A 1 24-JAN-2002;
Curagen Corporation (US)
Location/Qualifiers
1. .3741
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
/note="NOVX nucleic acid with homology to NOPE-like"

AX418406
Sequence 1 from Patent WO0206329.
AX418406
AX418406.1 GI:21523336
unidentified
unidentified
unclassified.
1
Rastelli, L., Shimkets, R.A., Zerhusen, B., Malyankar, U.M. and
Padigaru, M.
Human polynucleotides and polypeptides encoded thereby
Patent: WO 0206329-A 1 24-JAN-2002;
Curagen Corporation (US)
Location/Qualifiers
1. .3741
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
/note="NOVX nucleic acid with homology to NOPE-like"


```
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11:349H15"
/clone_lib="RPC1-11 Human Male BAC"
1..1477
/note="assembly_fragment"
1578..2785
/note="assembly_fragment"
2886..4859
/note="assembly_fragment"
4960..6416
/note="assembly_fragment"
6517..8739
/note="assembly_fragment"
8840..10868
/note="assembly_fragment"
10969..13737
/note="assembly_fragment"
13838..15329
/note="assembly_fragment"
15430..17938
/note="assembly_fragment"
18039..19981
/note="assembly_fragment"
20082..23587
/note="assembly_fragment"
23688..27447
/note="assembly_fragment"
27548..30181
/note="assembly_fragment
clone end:17
vector side:left"
30262..34717
/note="assembly_fragment"
34838..40101
/note="assembly_fragment"
40202..44039
/note="assembly_fragment"
44140..49950
/note="assembly_fragment"
50051..55555
/note="assembly_fragment"
55656..60496
/note="assembly_fragment
clone end:SP6
vector side:right"
60597..68544
/note="assembly_fragment"
68645..76940
/note="assembly_fragment"
77041..85480
/note="assembly_fragment"
85581..93905
/note="assembly_fragment"
94006..108063
/note="assembly_fragment"
108164..123496
/note="assembly_fragment"
123597..139965
/note="assembly_fragment"
140066..176282
/note="assembly_fragment"
/notes="assembly_fragment"
BASE COUNT 44272 a 43276 c 42765 g 43361 t 2608 others
ORIGIN
```

```
Query Match 13.8% Score 854.8; DB 2; Length 176282;
Best Local Similarity 64.5% Pred. No. 1.9e-193;
Matches 1930; Conservative 0; Mismatches 847; Indels 217; Gaps 35;
QY 3341 AGGCAACGGGAGAAAGACCGTCCCGAGCTGCAGGAATCAGGTGGAGCTGAGGTCA 3400
Db 154550 AGGGCAATGGGAGGAGAGAGTACACCCCGAGCTGCAGGAACGAGGTGGAGCTGAGTCA 154491
```

```
QY 3401 TTGTCCACTCCGACTTCGGTGCATCCAAAGATGTCCTGAGCTCCACCTCCCAAGACCTGG 3460
Db 154490 TTGTCCACTTCGACTTTAGTGCATCTAAACGGGAAACCTGAGACCTCCATCTCCAAGACCTGG 154431
QY 3461 AGCCAGAGAAACCACTGACTCGCAGAGACTCTGCCCTTCACGCTCTGAGCTGTGGATCTGT 3520
Db 154430 AGCCTGAGGAGACCCCTGCCTCCAGAGGCTCTGATCTCATCTCTGGGTGTGGGATCCAG 154371
QY 3521 CTCAGAGGAGACACTGGCTGGGACGGAGCTGGGAGGGTGCCAAACCAACACAGTGGGC 3580
Db 154370 GGCAGGGGCGACCTGGCTGGCAGAGGAGTTGGGAGGGTGTGAGCTGGCAGGCCCGGGGC 154311
QY 3581 CAGAGAGCTCACTGCTTGGCAGAAAGCAGCAGTCTCTGCTCTGCTCTGCTCTGCTCTGCT 3640
Db 154310 CAGACAGACTTACCTGCTTGCAGAGGACGAGTCTCTGCTCTGCTCTGCTCTGCTCTGCT 154251
QY 3641 AGCCAGCACTGCTATAGAGGAGGCCCTGGGAAAGCTGCCAGGCCCAAGCCCTGTGTC 3700
Db 154250 AGCCAGGCGAGGTGCTAGAGGAGACCCCTGGAGATAGTGTGCCAGCTCAAAATCCCTGCC 154191
QY 3701 CTCTAAACAGTCAAGCCAAAGCCTTCCAGGGGCCCTGTCTCTCTGCTCAGGTCCCTGTAG 3760
Db 154190 CTCTAGGAGCCAGCCAGGCCCTGCCAGATCCCCGGTCTCTCTCTCTGCTCTGCTCTGCT 154131
QY 3761 CAGA-----AGGCAGATATGCTCAGGAA-----CATGCCA 3791
Db 154130 CAGAGGATGTGGTTTGGGGCAGGAGGTATGATCACATAGGATCGATACCTGTGGCGG 154071
QY 3792 TGCATGGCTACAGTGTGTACTAGAGATATCCATAGTCTCTGAGGCTCTTGTAGGCTC 3851
Db 154070 TGTATGTCCATGTGTGCTGTAGATACATCATCAAGCCCTTTGGAGCTCTCTAAGTTG 154011
QY 3852 CTTTGGCTGGGTTTGGGGAGAA-----CTTTACTCTCCCTCATATTCGATCATCATAC 3905
Db 154010 CTTTGGCTGAGGGGAGAGAGAAACATGGATATTACATCCCCCTACTCTTTTGTGATAC 153951
QY 3906 AGGAGGGACTTGAGACACAGCTCTGTGTAAATGGACACGTGTGAAGTCGTGTGTGTGTG 3965
Db 153950 ACATGTGACATGTGAAAGACATACAGACATAGCTACATGTGATGTGCACATGTGTGAAG 153891
QY 3966 TGTGTGTGTG-----TGCTGTTGAGCTAGGAAACCTCTCCCTATGTGTAGTACTACTGTG 4020
Db 153890 TGCATGTATGCGTACTGGTGTGTGAGCTGGGAAACCGTGGCCCGAGCAGTGTGCTACTA 153831
QY 4021 GCCTAGTTGACCTCCGTGGCAGGATGGTG--TAACAGTGATCAGT--GCAGCTCTTTCA 4077
Db 153830 GCCTGATTTGCTCTCCAGTCCAGAACGGTGGCCCCACAGTGTGTCAGTCCCGCCCTGTGG 153771
QY 4078 GCTTTTAGCTTTGTCCACTAGCTTAGCTTTTATT--ACACTCTGAGAGTGTCTCCAGTGTGTG 4135
Db 153770 GCGCCCACTCCATCGCCCGCAGCTTTTATTACACTCTGAGAGTGTCTCCATGTGCTGT 153711
QY 4136 TCTACAAAGACAGCCCGCAGGCCCTCTCTGTGTGAGCTGTGTGAGAGAGTGCAGGTCAAC 4195
Db 153710 CTGACAAAGACAGTCCAGCCCAATTTCTC--CTGTCTGGCTGGTGTGGGTGCA----- 153660
QY 4196 TCCAGGGCCTATGACACCGGAGCCTACACAGCATGGCTGTGTCATCCCTGGCCTCTCTA 4255
Db 153659 -----GCAGGCTCTGAATGCTGGCATTTTCAGTGCATCATCCTCC 153619
QY 4256 AGGTCCAGATGTCTGGGTGAACCCAGCTCAGCTCCCTCTCTCTTTGAGCATCTCTGTAGC 4315
Db 153618 AGTCTCTTATTGCCCAATAGAGAGGTGGCCCTGGCTCCCTCCGAGGCACTCTGCAATT 153559
QY 4316 TAAATTTGTAAATCTGGGAAGTGCCTGGTTTGGGAAATCTTTTTCGCACTCTGTCCCTCT 4375
Db 153558 TAAATTTGTAAATCTGGGAAGTGCCTGGTTTGGGAAATCGCTTTCTCTCACTCTTCCCT 153499
QY 4376 CTGCCCCCTCTCTCAATTTGTTCTGTGTGATCTGTCT--CATGTCTCTTGTCTGATTC--C 4433
Db 153498 CTTCTCTGGCCCTGGCTGCTGTGTGTCTGTCTCTCCAGTCACTCTGCTCTCCAGCAC 153439
```


Perreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Glende, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lechoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McSwan, P., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodores, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (23-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
3 (bases 1 to 198295)

AUTHORS
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, A., Cooke, P., DeKrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Glende, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lechoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McSwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodores, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (14-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
4 (bases 1 to 198295)

AUTHORS
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, A., Cooke, P., DeKrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Glende, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lechoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McSwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodores, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (26-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
On Apr 26, 2002 this sequence version replaced gi:20148056.

All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L22971
Center clone name: 349_G_13

FEATURES
Source Location/Qualifiers
1. 198295
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-349G13"
complement(1..79)
/rpt_family="AluS"
complement(183..494)
/rpt_family="AluSg"
complement(1216..1416)
/rpt_family="L1MB8"
complement(1417..1709)
/rpt_family="AluSg"
complement(1710..2115)
/rpt_family="L1MB8"
2116..2169
/rpt_family="(TTTAA)n"
2186..2234
/rpt_family="AluSg/x"
complement(2236..2512)
/rpt_family="AluJb"
complement(2514..2631)
/rpt_family="L1MB8"
complement(2632..2941)
/rpt_family="AluY"
complement(2942..3104)
/rpt_family="L1MB8"
3181..3464
/rpt_family="AluSx"
3465..3711
/rpt_family="AluJb"
3712..3788
/rpt_family="(TA)n"
complement(4661..4940)
/rpt_family="AluJb"
complement(4946..5106)
/rpt_family="FRAM"
complement(5878..6193)
/rpt_family="AluYb8"
7197..7490
/rpt_family="AluSx"
7925..8226
/rpt_family="AluY"
complement(8227..8312)
/rpt_family="L1MB3"
8313..8347
/rpt_family="AT_rich"
complement(8350..8617)
/rpt_family="AluSx"
complement(8621..9274)
/rpt_family="L1MB3"
9403..9445
/rpt_family="AT_rich"
9706..10007
/rpt_family="AluJb"
complement(10940..11000)
/rpt_family="L1MB5"

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 17:04:36 ; Search time 14061.8 Seconds
(without alignments)
10674.595 Million cell updates/sec

Title: US-09-754-997A-1
Perfect score: 6176
Sequence: 1 atggcgggcggaacggg.....taataaattttccagac 6176

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*			
1:	em_estba:*		
2:	em_esthum:*		
3:	em_estin:*		
4:	em_estmu:*		
5:	em_estov:*		
6:	em_estpl:*		
7:	em_estro:*		
8:	em_htc:*		
9:	gb_est1:*		
10:	gb_est2:*		
11:	gb_hcc:*		
12:	gb_est3:*		
13:	gb_est4:*		
14:	gb_est5:*		
15:	em_estfun:*		
16:	em_estom:*		
17:	em_gss_hum:*		
18:	em_gss_inv:*		
19:	em_gss_pln:*		
20:	em_gss_vrt:*		
21:	em_gss_fun:*		
22:	em_gss_mam:*		
23:	em_gss_mus:*		
24:	em_gss_pro:*		
25:	em_gss_rod:*		
26:	em_gss_pbg:*		
27:	em_gss_vrl:*		
28:	gb_gss1:*		
29:	gb_gss2:*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1479	23.9	2626	11	AK079073 Mus muscu
2	883	14.3	947	11	AK020142 Mus muscu
3	805.8	13.0	849	14	CD352600 UI-M-GIO-
4	792.2	12.8	817	14	CA317106 UI-M-FW0-

5	737.6	11.9	748	14	CD348071
6	726.6	11.8	741	14	CD352400
7	720.6	11.7	944	13	BU840589
8	714.8	11.6	808	13	BU705518
9	706.4	11.4	719	14	CA751384
10	703.8	11.4	740	13	BU701299
11	690	11.2	713	14	CB524810
12	685.8	11.1	699	14	CD351162
13	639.6	10.4	900	12	BI185886
14	629.4	10.2	631	14	CB518504
15	618.2	10.0	671	14	BY719391
16	617.2	10.0	624	13	BU056331
17	599.8	9.7	651	10	BB624536
18	582	9.4	591	13	BU057120
19	581.6	9.4	637	10	BB665894
20	580	9.4	630	10	BB626002
21	574	9.3	621	9	AW049847
22	573	9.3	660	10	BB071741
23	561	9.1	583	14	CA895918
24	560.4	9.1	632	10	BB649997
25	556.4	9.0	759	14	CB988160
26	555.4	9.0	579	28	AZ746941
27	548.8	8.9	553	14	CA316452
28	546.2	8.8	703	14	BY712424
29	541.6	8.8	584	28	AZ745366
30	538.4	8.7	1137	11	AK082162
31	534.4	8.7	544	14	CA892088
32	529	8.6	551	13	BU705885
33	513.2	8.3	578	14	CB607425
34	509	8.2	509	14	CA873325
35	508.8	8.2	723	14	CA510263
36	505	8.2	528	4	BX519858
37	503.6	8.2	717	13	BX100607
38	502	8.1	565	14	CA883712
39	501.6	8.1	542	4	BX523249
40	500.8	8.1	504	14	CA876082
41	491.6	8.0	553	14	CA871246
42	485.6	7.9	528	13	BU580111
43	462.4	7.5	809	29	BZ222348
44	453.4	7.3	455	14	CA883713
45	452	7.3	533	10	BB681084

ALIGNMENTS

RESULT 1	AK079073	2626 bp	mRNA	linear	HTC 05-DEC-2002
LOCUS	Mus musculus adult male diencephalon cDNA, RIKEN full-length				
DEFINITION	enriched library, clone:9330155G14				product:hypothetical protein, full insert sequence.
ACCESSION	AK079073				
VERSION	AK079073.1				GI:26098261
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE					
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				

Db 795 AGTGTAGTGGTGGGCTCTGCTGACCCACCCCTTTTGTGCTCTGGTCCGACAGGA 854
 Qy 840 TGGAAAGCTATCTCCAGGATGTCTATGTTCTGGGCGGACCAATCTACTCATCGCCAG 899
 Db 855 TGGAAAGCTATCTCCAGGATGTCTATGTTCTGGGCGGACCAATCTACTCATCGCCAG 914
 Qy 900 CGCGAGCTCTGGCACTCTGGAGTCT-ATGTTCTGGCGAGCCAAAGCCCTCTCAGCGGTG 958
 Db 915 CGCCAGGCTCGGCACCTTTTGAATTTAAATTTTCCCAACCCACCAACCCCGCCAGCGCTT 974
 Qy 959 ACTTCGCCACCTCGGCGCTCTGAGCTCGAGTGTCTGCTGCCCCCAGCCATCTCGAGG--C 1016
 Db 975 AATTTTCCCCCTTCCGCTTCTTAAGTTTCAATTTCTTTTCCGCCACCCCTTTTCCAGGCC 1034
 Qy 1017 ACCGAGGCGCTCTCGCGGACCGCGGCGCAGCACCGCGCTTCTGTTGCGCGGCGCTCGG 1076
 Db 1035 GCCCAAGCCCTTTTCCGAAACCCGCGCCACCCCGCCCGTTCCTTTCCGCGCTTCGG 1094
 Qy 1077 GGAGCCACGCGCGCGCTGCACTGCTGTCGACGACGCGATCCGTTTGCAGCCCAATGGCGG 1136
 Db 1095 GGACCCCGCGCGCGCTGCACTGCTGTCGACGACGCGAATCCGTTTGCAGCCCAATGGCGG 1154
 Qy 1137 CGTCAAGTGCAGGCGCGTGGCGGAGCTTGTGTCATCTCATGATCGCGCTTCAGGACGC 1196
 Db 1155 CGTCAAGTGCAGGCGCGTGGCGGAGCTTGTGTCATCTCATGATCGCGCTTCAGGACGC 1214
 Qy 1197 TGGCTACTACGAGTGCCTAGCAGAAACAGCGCGGAACTGCTGTGCGCTGCGCCCT 1256
 Db 1215 TGGCTACTACGAGTGCCTAGCAGAAACAGCGCGGAACTGCTGTGCGCTGCGCCCT 1274
 Qy 1257 GCGCGTAGTGTGCGCGAGGCGCTGCCAGCGCCCGACTCGGGTCACAGCCACCGCGCT 1316
 Db 1275 GCGCGTAGTGTGCGCGAGGCGCTGCCAGCGCCCGACTCGGGTCACAGCCACCGCGCT 1334
 Qy 1317 GAGCAGCTCTCTGTGCTGTGCTGAGGAGCGGCGCTGAG-TTGCACAGCGAGCAATCA 1375
 Db 1335 GAGCAGCTCTCTGTGCTGTGCTGAGGAGCGGCGCTGAGTTTGCACAGCGAGCAATCA 1394
 Qy 1376 TTGGCTT-CTCTCTTCTACTACAAAGCAAGGCGGAGTGGAATG--TGAGTACAG 1431
 Db 1395 TTGGCTTCTCTCTTCTACTACAAAGCAAGGCGGAGTGGAATGTTGGAGTACAG 1454
 Qy 1432 TTTCAGTAAACAAATGACACACAG------GCTGCAGGTTCG 1469
 Db 1455 TTTCAGTAAACAAATGGAACCAAGAGAGCTTGGAGGGTTTCGGGAACTCG 1514
 Qy 1470 GGACTTGAACCCAAACAGGATATGATTTCTACTGTTGGCTTACTCCAGCTGGGCGC 1529
 Db 1515 GAAACCCCAACCCCGATTAATGAAGTTTAAAGTGTGGCTTAACCTTCCAAAGCTGGG 1574
 Qy 1530 CAGCC-----GAACCTCCAGCCAG--CCCTGTGTCATACACTGGAGCATGTCCCGAG 1580
 Db 1575 GGGCCAGGCGGAAACCTTCAAGCCAGGCGCTGCTGTCATACACTGGAGCATGTCCCGAG 1634
 Qy 1581 CGCAGCACCCAGCTTACTTCTGACGCCCCCAACCCCTCGGCATC-AGGGTGGCATGGC 1639
 Db 1635 CGCACACCCAGCTTACTTCTGACGCCCCCAACCCCTCGGCATC-AGGGTGGCATGGC 1694
 Qy 1640 TGCCCCCTGCTCCAGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1699
 Db 1695 TGCCCCCTGCTCCAGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1754
 Qy 1700 TGGGAAG---GAAATGAGTTTCTCAGCGAGTGTCTGGAATGAGACACAACTTA 1756
 Db 1755 TGGGAAGGAAAGATGAGTTTCTCAGCGAGTGTCTGGAATGAGACACAACTTA 1814
 Qy 1757 GCTTAACTCACTTCAGCAGCAACAGTGTACCGAGTCCGAGTTCAGTCCGACTGGCG 1816
 Db 1815 GCTTAACTCACTTCAGCAGCAACAGTGTACCGAGTCCGAGTTCAGTCCGACTGGCG 1874
 Qy 1817 CT-GGCTATGGAGTCCCTTCTCAGTGGATGACAGCAGCAGCTGGTGTGACACACAG 1875

Db 1875 CTGGGCTATGGAGTCCCTTCTCAGTGGATGACAGCAGCAGCAGCAGCAGCAGCAGCAG 1934
 Qy 1876 AGCCATGTT 1884
 Db 1935 AGCCATGTT 1943
 RESULT 2
 AKO20142
 LOCUS
 DEFINITION
 Mus musculus 12 days embryo male mrna linear HTC 05-DEC-2002
 region cDNA, RIKEN full-length enriched library, clone:6720465A10
 product:hypothetical protein, full insert sequence.
 AKO20142
 AKO20142.1 GI:12860634
 VERSION
 HTC; CAP trapper.
 KEYWORDS
 Mus musculus (house mouse)
 SOURCE
 Mus musculus
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci, P. and Hayashizaki, Y.
 TITLE
 High-efficiency full-length cDNA cloning
 JOURNAL
 Mech. Enzymol. 303, 19-44 (1999)
 MEDLINE
 99279253
 PUBMED
 10349636
 REFERENCE
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 JOURNAL
 20499374
 MEDLINE
 11042159
 PUBMED
 REFERENCE
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multipillarary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 JOURNAL
 20530913
 MEDLINE
 11076861
 PUBMED
 REFERENCE
 4
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,
 Quackenbush, J., Schriml, L.M., Staab, F., Suzuki, R., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
 Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Buit, C.,
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
 Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombaeerts, P., Nordone, P.,
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
 Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L.,
 Wyshaw-Borja, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
 and Hayashizaki, Y.
 TITLE
 Functional annotation of a full-length mouse cDNA collection
 JOURNAL
 Nature 409 (6821), 685-690 (2001)
 MEDLINE
 21085860
 PUBMED
 11217851
 REFERENCE
 5
 The FANTOM Consortium and the RIKEN Genome Exploration Research

Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mouseefi.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

FEATURES

Location/Qualifiers

1. 849

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE: 6856471"

/tissue_type="whole brain"

/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH BMAP G10"

/note="Organ: Brain; Vector: pYX-Asc; Site 1: Ecor I;

Site 2: Not I; The library was constructed according

Bonaldi, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

with Ecor I adaptor, digested with NotI and then cloned

directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

is ACGGAGACAG. This library was created for the University

Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the

Developing Mouse Nervous System', supported by National

Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,

program coordinator." 237 t 7 others

BASE COUNT 176 a 225 c 204 g 237 t 7 others

ORIGIN

Query Match 13.0%; Score 805.8; DB 14; Length 849;

Best Local Similarity 98.6%; Pred. No. 3.5e-125;

Matches 840; Conservative 0; Mismatches 9; Indels 3; Gaps 3;

QY 4390 ATTGTTCTGGTGATCTGTCTCATGTCATCTTGTCTGATTAFCCTGGGGCCCTTCTCTTT 4449

DB 1 ATTTGTTCTGGTGATCTGTCTCATGTCATCTTGTCTGATTAFCCTGGGGCCCTTCTCTTT 60

QY 4450 CCCATGATGCCCTGATTTCTCTCACTGCTGTTTTCATTTCTGTCGCCATGCTTGTCTTT 4509

DB 61 CCCATGATGCCCTGATTTCTCTCACTGCTGTTTTCATTTCTGTCGCCATGCTTGTCTTT 120

QY 4510 ATGTCGTGTTTCTCGTCCCTGAGTTCACCTATGCACTTCCCTTAACACATGACTA 4569

DB 121 ATGTCGTGTTTCTCGTCCCTGAGTTCACCTATGCACTTCCCTTAACACATGACTA 180

QY 4570 CCTCATGTCGTCTCAGACCATAGTGTGACCCCTGGGTCCCAACAGCTCCCTGCAACC 4629

DB 181 CCTCATGTCGTCTCAGACCATAGTGTGACCCCTGGGTCCCAACAGCTCCCTGCAACC 240

QY 4630 GCTTTCTGGGAGATGAGCCCACTCCAAGTAGATCTGGAAGAAGCCCTTGTGGCTTGTTC 4689

DB 241 GCTTTCTGGGAGATGAGCCCACTCCAAGTAGATCTGGAAGAAGCCCTTGTGGCTTGTTC 298

QY 4690 TGGCTGCCCTCCCTGGTGTGATGAGAGGTTTCTATGGAAGATGAGTCCAGG 4749

DB 299 TGGCTGCCCTCCCTGGTGTGATGAGAGGTTTCTATGGAAGATGAGTCCAGG 358

QY 4750 CTGCACAGGGGAACCCCAAGAGGGGTAGGAGTGAACCAAGAGGCTGAAAAAATG 4809

DB 359 CTGCACAGGGGAACCCCAAGAGGGGTAGGAGTGAACCAAGAGGCTGAAAAAATG 418

QY 4810 GCTGCCACCCATCTGCACAGAGATGGGTGTGTCTTTTGTACGTGCAAGTCTGGCTGAA 4869

DB 419 GCTGCCACCCATCTGCACAGAGATGGGTGTGTCTTTTGTACGTGCAAGTCTGGCTGAA 478

QY 4870 ACTGAAGGGGTGAGGAGAGGGAGCTACTTGGGGCTGCCATGGCTCAGTTCCTGACCCCTG 4929

DB 479 ACTGAAGGGGTGAGGAGAGGGAGCTACTTGGGGCTGCCATGGCTCAGTTCCTGACCCCTG 538

QY 4930 GAGCCCTGAACCTGGCTTCAGAGTAGCAAGAGGTTTCTCCAGATGCTGTGAAGGAAATG 4989

DB 539 GAGCCCTGAACCTGGCTTCAGAGTAGCAAGAGGTTTCTCCAGATGCTGTGAAGGAAATG 598

QY 4990 CTTTGTGATAGGAAAAAGGGCGGCTGCTCATTTATTTATTTTATCTTTTACACTGAATCCC 5049

DB 599 CTTTGTGATAGGAAAAAGGGCGGCTGCTCATTTATTTATTTTATCTTTTACACTGAATCCC 658

QY 5050 AAAATCATCTTACCACAAAGGGCCAGCTGACTGTGTAATTTCTTGAGTCAACAAGAGCCAT 5109

DB 659 ANAATCATCTTACCACAAAGGGCCAGCTGACTGTGTAATTTCTTGAGTCAACAAGAGCCAT 718

QY 5110 GCCATCTCTCTGGTTTCTCACCTCAGTCATGTCAGAGTAATTTGTGAGTCCAGTGGCATCT 5169

DB 719 GCCATCTCTCTGGTTTCTCACCTCAGTCATGTCAGAGTAATTTGTGAGTCCAGTGGCATCT 778

QY 5170 GTGCTCTTGTGTCACATCTTCTTATTTTCAACTGGCTGSCACATCAAGTGTAACTCTGGC 5229

DB 779 GTGCTCTTGTGTCACATCTTCTTATTTTCAACTGGCT-GCATCAAGTGTAACTCTGGC 837

QY 5230 TTCTGGGCCAAG 5241

DB 838 TTCTGGGCCAAG 849

RESULT 4

CA317106

LOCUS

DEFINITION

CA317106

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CA317106 817 bp mRNA linear EST 26-NOV-2002
UI-M-FW0-cbm-d-21-0-UI-r1 NIH BMAP_FW0 Mus musculus cDNA clone
IMAGE: 6811462 5', mRNA sequence.

CA317106

EST.

CA317106.1 GI:24535230

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1. 817

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE: 6811462"

/tissue_type="whole brain"

/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH BMAP_FW0"

/note="Organ: Brain; Vector: pYX-Asc; Site 1: Ecor I;

Site 2: Not I; The library was constructed according

Bonaldi, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel.

First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

with Ecor I adaptor, digested with NotI and then cloned

directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

is ACGGAGACAG. This library was created for the University

Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the

gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 180 a 254 c 234 g 148 t 1 others
ORIGIN

Query Match 12.8%; Score 792.2; DB 14; Length 817;

Best Local Similarity 98.8%; Pred. No. 6.8e-123;

Matches 808; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

Qy 3009 AGCTCGGCTGGGGCTCCAGTGTCCCTGCTGCCCATGAGTTGGAGTCCCTCGTGCATCC 3068
Db 1 AGCTCGACTTGGGGCCCCCAGTGTCCCTGCTGCCCATGAGTTGGAGTCCCTCGTGCATCC 60
Qy 3069 TCGTCCCGAGATTGGTCCACCCAGTGTCCAGATGTCGAGTGTGGAACAAGGCTGAAGTACACAG 3128
Db 61 TCGTCCCGAGATTGGTCCACCCAGTGTCCAGATGTCGAGTGTGGAACAAGGCTGAAGTACACAG 120
Qy 3129 CTTTATGGTGGCAGTGTTCAGATTGCGGGGCCACTCCAAAGAGAAAGATCTCTCGGC 3188
Db 121 CTTTATGGTGGCAGTGTTCAGATTGCGGGGCCACTCCAAAGAGAAAGATCTCTCGGC 180
Qy 3189 TCAGGCGAGGGGACAAATCGGGCAGGCTCTGGGAGGCTGTGAGTGTGCCACAGGGTAG 3248
Db 181 TCAGGCGAGGGGACAAATCGGGCAGGCTCTGGGAGGCTGTGAGTGTGCCACAGGGTAG 240
Qy 3249 TGGTCCAGGCGGCTCTGACCGGTCTGCTGCTCCAGCGGACCGGCGAGACACT 3308
Db 241 TGGTCCAGGCGGCTCTGACCGGTCTGCTGCTCCAGCGGACCGGCGAGACACT 300
Qy 3309 GCTGCTCAAGCCCTGTGTATGCGGCATAAAGAGCAAGCGGAGAAAGCGTCCCC 3368
Db 301 GCTGCTCAAGCCCTGTGTATGCGGCATAAAGAGCAAGCGGAGAAAGCGTCCCC 360
Qy 3369 AGCTCGAGGAATCAGTGGAGTGAAGTCAATGTTCATCTCCAGTTCGGTGCATCAA 3428
Db 361 AGCTCGAGGAATCAGTGGAGTGAAGTCAATGTTCATCTCCAGTTCGGTGCATCAA 420
Qy 3429 AGGATGCTGACCTCCAGCTCCAGACCTGGAGCCAGAGAACCACTGACTGCAGAGAC 3488
Db 421 AGGATGCTGACCTCCAGCTCCAGACCTGGAGCCAGAGAACCACTGACTGCAGAGAC 480
Qy 3489 TCTGCTTCCAGCTCTGAGCTGTGGATCTGTCTCAAGGAGCAGACTGGCTGGGACGGGA 3548
Db 481 TCTGCTTCCAGCTCTGAGCTGTGGATCTGTCTCAAGGAGCAGACTGGCTGGGACGGGA 540
Qy 3549 GCTGGGAGGTGCAACCAACAGTGGGCCAGAGAGGCTCACTGCTTGCAGAGC 3608
Db 541 GCTGGGAGGTGCAACCAACAGTGGGCCAGAGAGGCTCACTGCTTGCAGAGC 600
Qy 3609 AGCCAGTGCCTCTGCTCTGCTCAGACCTCCAGCCAGCACTCTATAGAGGAGGCC 3668
Db 601 AGCCAGTGCCTCTGCTCTGCTCAGACCTCCAGCCAGCACTCTATAGAGGAGGCC 660
Qy 3669 TGGGAAAGCTGCGCAGGCCCAAGCCCTGTGTCTTAACTGTCAGCCCAAGCCCTTCCAG 3728
Db 661 TGGGAAAGCTGCGCAGGCCCAAGCCCTGTGTCTTAACTGTCAGCCCAAG-CTTCCAG 719
Qy 3729 GGCCCTCTGTCTCTGCTCAGGTCCTCGTGAAGAGGAGGAGATATGCTCAGGAACATG 3788
Db 720 NGCCCTCTGTCTCTGCTCAGGTCCTCGTGAAGAGGAGGAGATATGCTCAGGAACATG 779
Qy 3789 CCATGATGGCTACACATGTGTGTACTAGATATCCA 3826
Db 780 CCATGATGGCTACACATGTGTGTACTAGATATCCA 817

RESULT 5

CD348071

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CD348071 748 bp mRNA linear EST 29-MAY-2003
UI-H-FYO-cfm-h-04-0-UI.r1 NIH_BMAP_FYO Mus musculus cDNA clone
IMAGE: 6851477 5', mRNA sequence.
CD348071
EST.
CD348071.1 GI:31139586
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 748)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs-rc@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library prepared by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
The following repetitive elements were found in this cDNA
sequence: 179-203, >(CA)n#Simple_repeat (matched complement)
Seq primer: pYX-5.

Location/Qualifiers

1..748
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6851477"
/issue_type="whole brain"
/cissue_type="embryo 13.5, 14.5, 16.5, 17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FYO"
/notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

BASE COUNT 129 a 202 c 177 g 236 t 4 others

ORIGIN

Query Match 11.9%; Score 737.6; DB 14; Length 748;
Best Local Similarity 98.9%; Pred. No. 1e-113;
Matches 740; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 3775 GGCTCAGGACATGCCATCGCTACACATGTGTGTACTAGATATCCATAGTCCT 3834
Db 1 GGCTCAGGACATGCCATCGCTACACATGTGTGTACTAGATATCCATAGTCCT 60
Qy 3835 TGGAGCTCTTAGGGTCTTTGGCTGGGGTGGGAGAACTTTACTCTCCCTCATATCT 3894
Db 61 TGGAGCTCTTAGGGTCTTTGGCTGGGGTGGGAGAACTTTACTCTCCCTCATATCT 120
Qy 3895 GCATCATACAGAGGGAGCTTGAGACACAGCTCTGTGTAAAGGACACGCTGTGAAGTCGT 3954


```

Db      601  CTCTGGAGTCTATGTCTGCGAGGCCAACAGCCCGCCGACCGCGTGAATTTGCGCCACTCGGCG 660
Qy      975  TGCTGAGCTCCGAGTCTGCTGCTGCCAGCAGCATCTGCGAGCCACCGAGCGCTCTCGCG 1034
Db      661  TGCTGAGCTCCGAGTCTGCTGCTGCCAGCAGCATCTGCGAGCCCGCCGAGCGCTCTCGCG 720
Qy      1035 GACGCGGCGCCAGCACCGCGCG 1055
Db      721  GACGCGGCGCCAGCACCGCGCG 741

RESULT 7
BUB40589
LOCUS
DEFINITION BUB40589 944 bp mRNA linear EST 16-OCT-2002
AGENCOURT 8946353 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6490299
5', mRNA sequence.
ACCESSION BUB40589
VERSION BUB40589.1 GI:24024989
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
JOURNAL Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
COMMENT 1. (bases 1 to 944)
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM14039 Row: 0 Column: 04
High quality sequence stop: 535.
FEATURES
source
1..944
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6490299"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_94"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 195 a 270 c 227 g 252 t
ORIGIN
Query Match 11.7%; Score 720.6; DB 13; Length 944;
Best Local Similarity 96.1%; Pred. No. 6.8e-111;
Matches 793; Conservative 0; Mismatches 24; Indels 8; Gaps 5;

Qy      4254 TAAGGTCCAGATGTCGGGTGAACCCAGCTCAGCTCCCTCTCTTTGAGCATCTGTGA 4313
Db      1 TAAGGTCCAGATGTCGGGTGAACCCAGCTCAGCTCCCTCTCTTTGAGCATCTGTGA 60
Qy      4314 CCTAATTTGTAATCTGGGAGTCCCTGGTTGGGAATCTTCTTGGACCCCTGTCCT 4373
Db      61 CCTAATTTGTAATCTGGGAGTCCCTGGTTGGGAATCTTCTTGGACCCCTGTCCT 120
Qy      4374 CTCTGCCCTTCTCTCATTTGTTCTGTTGATCTGCTCATGTCATCTTGTGCAATATCC 4433
Db      121 CTCTGCCCTTCTCTCATTTGTTCTGTTGATCTGCTCATGTCATCTTGTGCAATATCC 180
Qy      4434 TGGGCGCCCTTCTCTCTTCCCATGATGCCCTGATTTCCCTACTGCTGTTTCTATCTGTC 4493

```

```

Db      181  TGGGGCCCTTCTCTTTTCCCATGATGCCCCCTGATTTCTCCTCACTGCTGTTTTTCAATTTCTGTC 240
Qy      4494  TGCCATGCTTGTCTTTATGTCGTGTGTTTCTGCTCCCTGAGTTCAACCTATGCAACCCCTTT 4553
Db      241  TGCCATGCTTGTCTTTATGTCGTGTGTTTCTGCTCCCTGAGTTCAACCTATGCAACCCCTTT 300
Qy      4554  CTTAAACAACATGACTACCTCATGTCTGCTTTCAGACCATAGTGTGACCCCTGGTCCCAAC 4613
Db      301  CCTAAACAACATGACTACCTCATGTCTGCTTTCAGACCATAGTGTGACCCCTGGTCCCAAC 360
Qy      4614  AGCTCCCTGCGCAACCGGCTTCTGCGGAGATGAGCCCACTCCAGTAGATCTGGAAGAG 4673
Db      361  AGCTCCCTGCGCAACCGGCTTCTGCGGAGATGAG-CCCACTCCAGTAGATCTGGAAGAG 419
Qy      4674  ACCCTTGTGGCTTGTCTGCTGCTCCCTCCCTTGGTGTTCAGATGAGAGAGTTTCTATGG 4733
Db      420  A-CCTTGTGGCTTGTCTGCTGCTCCCTCCCTTGGTGTTCAGATGAGAGAGTTTCTATGG 478
Qy      4734  AAGAGATGATCCAGGCTGCA CAGGGGAACCCCAAGAGAGGAGTGGAGTGAACCAAG 4793
Db      479  AAGAGATGATCCAGGCTGCA CAGGGGAACCCCAAGAGAGGAGTGGAGTGAACCAAG 538
Qy      4794  AGCTGTAAGAAAATGGCTGCCACCATCTGCACAGAGAGATGGGTGTGCTTTGAGC 4853
Db      539  AGCTGTAAGAAAATGGCTGCCACCATCTGCACAGAGAGATGGGTGTGCTTTGAGC 598
Qy      4854  TGCAGTCTCTGCTGAAAATGGAAGGGGTGAGGAGAGGAGGAGTACTGGGGCTGCATGGCT 4913
Db      599  TGCAGTCTCTGCTGAAAATGGAAGGGGTGAGGAGAGGAGGAGTACTGGGGCTGCATGGCT 658
Qy      4914  CAGTTCCTGACCCCTGGAGCCCTGAAACCTGGCTTCAGAGTAGCAAGAGATTTCTTCAAG 4973
Db      659  CAGTTCCTGACCCCTGGAGCCCTGAAACCTGGCTTCAGAGTAGCAAGAGATTTCTTCAAG- 717
Qy      4974  ATGCTGTAGGAGAGTCTTTGATAGGAAAAGGGGGGCTGGCTCATTTTA--TTTATCT 5031
Db      718  ATGCTGTAGGAGAGGAGCTTTGATAGGAAAAGGGGGGCTGGGTCCATTTAATTTTAACT 777
Qy      5032  TTCTTTTACACTGAA--TCCCAAAATCATCTTTACCAAAAGGGGCC 5073
Db      778  TGCTTTACACCGGATCCCAAAATCATCTTTACCAAAAGGGGCC 822

BUB705518 808 bp mRNA linear EST 09-OCT-2002
UI-M-FOO-cac-e-03-0-UI.r1 NIH_BMAP_F00 Mus musculus cDNA clone
IMAGE: 6409322 5', mRNA sequence.
BUB705518
BUB705518.1 GI:23635078
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1. (bases 1 to 808)
NIH-MGC http://imgc.ncbi.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
cDNA Library Preparation: Dr. Jim Lin, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
The following repetitive elements were found in this cDNA
sequence: 30-59, >(TAAA)n#simple_repeat (matched complement)
Seq primer: pyX-5.

```

[illegible][illegible]

QY 673 GCCTCGCTCACTGTGGCCCTCAGAGGCTCTTTGAGGCTACACAGGGGCGAGGATGTGGTC 732
Db 1 GCCTCGCTCACTGTGGCCCTCAGAGGCTCTTTGAGGCTACACAGGGGCGAGGATGTGGTC 60
QY 733 ATTGTGGCAGCCCGCAGAGAAACACACGGTAGTGTCTGGACAGAAATGTAGTAGTGTC 792
Db 61 ATTGTGGCAGCCCGCAGAGAAACACACGGTAGTGTCTGGACAGAAATGTAGTAGTGTC 120
QY 793 GTGGCCCTCTGTGACCCACCCCTTTTGTCTCTGCTGGTCCGACAGAGATGGAAGCCCTATC 852
Db 121 GTGGCCCTCTGTGACCCACCCCTTTTGTCTCTGCTGGTCCGACAGAGATGGAAGCCCTATC 180
QY 853 TCACAGGATGTCTGTCTGTGGCCGAGCAACATCTACTCTATCGCAGCGCGAGCCCTCGG 912
Db 181 TCACAGGATGTCTGTCTGTGGCCGAGCAACATCTACTCTATCGCAGCGCGAGCCCTCGG 240
QY 913 CACTCTGAGTCTATGTCTGCGGAGCAACAAAGCCCTCTACGGTGTACTTGGCACTGGG 972
Db 241 CACTCTGAGTCTATGTCTGCGGAGCAACAAAGCCCGCAGCGCTGTATTTGGCACTGGG 300
QY 973 GCTGCTGAGTCTGCGGAGTCTGTCTGCGGAGCAACATCTGCGAGGACCCGAGGCGCTCTCG 1032
Db 301 GCTGCTGAGTCTGCGGAGTCTGTCTGCGGAGCAACATCTGCGAGGCGCGGAGGCGCTCTCG 360
QY 1033 CGGACGCGGCGCAGCACCGCGCTTCTGTGCGGCGCTTCCGGGAGCCACGCGCCCGCG 1092
Db 361 CGGACGCGGCGCAGCACCGCGCTTCTGTGCGGCGCTTCCGGGAGCCACGCGCCCGCG 420
QY 1093 CTGCACTGTGCTGCAAGCGGATCCCTGTGCGGAGCAACATCTGCGGCGCTTCAAGGTCAGGGC 1152
Db 421 CTGCACTGTGCTGCAAGCGGATCCCTGTGCGGAGCAACATCTGCGGCGCTTCAAGGTCAGGGC 480
QY 1153 GGTGGCGGAGCTGTGTCATCACTACATCGGCTGCGAGAGCGTGGCTACTACAGTGC 1212
Db 481 GGTGGCGGAGCTGTGTCATCACTACATCGGCTTCCAGAGCGTGGCTACTACAGTGC 540
QY 1213 GTAGCAGAAACAGCGCGGAGTCTGCTGTGCGGCTTCCGGGAGCCACGCGCCCGCG 1272
Db 541 GTAGCAGAAACAGCGCGGAGTCTGCTGTGCGGCTTCCGGGAGCCACGCGCCCGCG 600
QY 1273 GAGGGCTGCGCAGCGCCCGGACTCGGGTCAACAGCAGCGCGCTGAGCAGCTCTCTGTG 1332
Db 601 GAGGGCTGCGCAGCGCCCGGACTCGGGTCAACAGCAGCGCGCTGAGCAGCTCTCTGTG 660
QY 1333 CTGCTGCGCTGGAGCGGCTGAGTTCACAGCGAGCAATCATTTGCTCTCTCTCA 1391
Db 661 CTGCTGCGCTGGAGCGGCTGAGTTCACAGCGAGCAATCATTTGCTCTCTCTCTCA 719

RESULT 10
BU701299
LOCUS
DEFINITION
UI-M-EXO-bwy-i-18-0-UI.r1 NIH BMAP_EXO Mus musculus CDNA clone
IMAGE: 5704793 5', mRNA sequence.
ACCESSION
BU701299
VERSION
BU701299.1 GI:23624921
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 740)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: csapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

FEATURES

source

Seq primer: pYX-5,
Location/Qualifiers
1..740
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 5704793"
/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DH10B (T1 phage
/clone_lib="NIH BMAP_EXO"
/note="Organ: Brain; Vector: pYX-Asc; Site_1: EcoR I;
Site_2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is GTGCGTGGAA. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 173 a 225 c 200 g 142 t
ORIGIN

Query Match

Best Local Similarity 11.4%; Score 703.8; DB 13; Length 740;
Matches 738; Conservative 0; Mismatches 2; Indels 3; Gaps 3;

QY 1959 CCCCACCCAGATCTCTGGATACAACTCTACTGGGAGAGGTGGGAACAGAGGAGGCG 2018

Db 1 CCCCACCCAGATCTCTGGATACAACTCTACTGGGAGAGGTGGGAACAGAGGAGGCG 60

QY 2019 AGATGTGACCGCCCGCCAGGGGCTGCTGGAGATCAAGCTTGGGAGCTGGGCGCCGTGCG 2078

Db 61 AGATGTGACCGCCCGCCAGGGGCTGCTGGAGATCAAGCTTGGGAGCTGGGCGCCGTGCG 120

QY 2079 GCTGAAGAAGAAGTGAAGCAGTATGAACCTGACCCAGTTAGTCCCTGGGAGGCGCTACGA 2138

Db 121 GCTGAAGAAGAAGTGAAGCAGTATGAACCTGACCCAGTTAGTCCCTGGGAGGCGCTACGA 180

QY 2139 GGTGAAGCTCGTAGCTTTCACAAACACAGGAGCGGCTACGCTGTGTGGAAGGCGCA 2198

Db 181 GGTGAAGCTCGTAGCTTTCACAAACACAGGAGCGGCTACGCTGTGTGGAAGGCGCA 240

QY 2199 GACGGAAGAAGGCGCCACCGCCAGACCTGCTATCCAGAGGGGCGCCACCGCTGCTCTGC 2258

Db 241 GACGGAAGAAGGCGCCACCGCCAGACCTGCTATCCAGAGGGGCGCCACCGCTGCTCTGC 300

QY 2259 CCATGTCCACGAGAGTCAAAACAGCTCCACTTCCTCCATTTGGCTTGGTGAAGAAGCCAGA 2318

Db 301 CCATGTCCACGAGAGTCAAAACAGCTCCACTTCCTCCATTTGGCTTGGTGAAGAAGCCAGA 360

QY 2319 CTTTACACATGTCAAGATTGTCACTACATCTGTACGCTTGGGCGGCTCAGGAA 2378

Db 361 CTTTACACATGTCAAGATTGTCACTACATCTGTACGCTTGGGCGGCTCAGGAA 420

QY 2379 TGTCTCCCTGGTCCACCTACTATACAGCTCTCGAGAGAGCAATCTCTCATTTGGCGGCTGAA 2438

Db 421 TGTCTCCCTGGTCCACCTACTATACAGCTCTCGAGAGAGCAATCTCTCATTTGGCGGCTGAA 479

QY 2439 ACCATTTACCAAGTACGAGTTTGGCGGTACAGTCCCAAGGAGTGATATGATGGGCGCTT 2498

Db 480 ACCATTTACCAAGTACGAGTTTGGCGGTACAGTCCCAAGGAGTGATATGATGGGCGCTT 539

2499 TGGCTCCGCTGTAGAACCTCCACCCCTCCGCTGACCGGCTTCAACACCTCTCTTGACCT 2558
 Db TGGCTCCGCTGTAGAACCTCCACCCCTCCGCTGACCGGCTTCAACACCTCTCTTGACCT 599

2559 GCGCTGAGCCCTGTAGAACCTCCACCGTTGGTTACACTGGTGTCCCGCCACGAGCC 2618
 Db GCGCTGAGCCCTGTAGAACCTCCACCGTTGGTTACACTGGTGT-CCCCCAGGAGCC 658

2619 CAATGGTGTAGATTGGAGTATCTAATTTCTTACAGCAACACACCCAGCCCGAACA 2678
 Db CAATGGTGTAGATTGGAGTATCTAATTTCTTACAGCAACAA-CACACCCAGCCCGAACA 717

2679 CCAGTGGACATGCTCACACAG 2701
 Db CCAGTGGACATGCTCACACAG 740

RESULT 11
 CB524810
 LOCUS
 DEFINITION UI-M-FY0-cey-g-12-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
 IMAGE: 6843781 5', mRNA sequence.

ACCESSION CB524810
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Mus musculus (house mouse)

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 JOURNAL 1 (bases 1 to 713)
 COMMENT NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.uiowa.edu/distribution/mousefi.html
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.
 Location/Qualifiers
 1..713
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE: 6843781"
 /tissue_type="whole brain"
 /dev_stages="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_FY0"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is AGCAGAGACAG. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

141 a 235 c 185 g 148 t 4 others
 BASE COUNT
 ORIGIN

Query Match 11.2%; Score 690; DB 14; Length 713;
 Best Local Similarity 98.6%; Pred. No. 9.7e-106;
 Matches 704; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 2489 ATGGGCGCTTTGGCTCGTGTAGAACGCTCCACCCCTCCGCTGACCGGCTTCAACACCTC 2548
 Db 1 ATGGGCGCTTTGGCTCGTGTAGAACGCTCCACCCCTCCGCTGACCGGCTTCAACACCTC 59

QY 2549 CTTCTGACCTGCGCTGACGCCCCCTGACACCATCCACCGTTGGTTACACTGGTGTCCCGC 2608
 Db 60 CTTCTGACCTGCGCTGACGCCCCCTGACACCATCCACCGTTGGTTACACTGGTGTCCCGC 119

QY 2609 CCACGAGCCCAATGGTGTAGATTGTGGAGTATCTAATTTCTTACAGCAACACACACCC 2668
 Db 120 CCACGAGCCCAATGGTGTAGATTGTGGAGTATCTAATTTCTTACAGCAACACACACCC 179

QY 2669 AGCCCGAACACACGACATGCTCACCACAGAGGGAACATCTTCAAGTGCAGAGTGC 2728
 Db 180 AGCCCGAACACACGATGGACATGCTCACCACAGAGGGAACATCTTCAAGTGCAGAGTGC 239

QY 2729 ATGGCCCTAGAGAGTACACTCGGTATTTCTTCAAGATGGGAGCCCGCACAGAGTGGGGC 2788
 Db 240 ATGGCCCTAGAGAGTACACTCGGTATTTCTTCAAGATGGGAGCCCGCACAGAGTGGGGC 299

QY 2789 CTGGGCGCTTTTCCCGCTTGCAGGATGTGATTCTCTGCAAGAGACATTTCTCAGACTCCT 2848
 Db 300 CTGGGCGCTTTTCCCGCTTGCAGGATGTGATTCTCTGCAAGAGACATTTCTCAGACTCCT 359

QY 2849 TGGATGTGACCGCGCTCAGCGGCATCATCGTGGTGTCTGCTGGGCTTCTCTGCTCCTCC 2908
 Db 360 TGGATGTGACCGCGCTCAGCGGCATCATCGTGGTGTCTGCTGGGCTTCTCTGCTCCTCC 419

QY 2909 TGGCTCTCATGTGTGCTGGCTACGACAAAGCTCCACAGGAGAGCCCTTCCCGGATTGT 2968
 Db 420 TGGCTCTCATGTGTGCTGGCTACGACAAAGCTCCACAGGAGAGCCCTTCCCGGATTGT 479

QY 2969 CCTCTCAGGACACCCAGGAAACCCAGCGCTCTACACAGAGCTCGGCTGGGCTCCCA 3028
 Db 480 CCTCTCAGGACACCCAGGAAACCCAGCGCTCTACACAGAGCTCGGCTGGGCTCCCA 539

QY 3029 GTGTCCCTGTGCCCATAGTGTGAGTGCCTCTGTGCAATCTCTGCTCCCGAGATTGTGCC 3088
 Db 540 GTGTCCCTGTGCCCATAGTGTGAGTGCCTCTGTGCAATCTCTGCTCCCGAGATTGTGCC 599

QY 3089 CACCACCTCAGATGTGGAAGACAGCTGAGTACAGACGCTTATGGTGGCAGTGT 3148
 Db 600 CACCACCTCAGATGTGGAAGACAGCTGAGTACAGACGCTTATGGTGGCAGTGT 659

QY 3149 CAGATTCCGGGGCCACTCCAAAGAGAAAGATCTCTGGGCTCAGGAGGGGAC 3202
 Db 660 CAGATTCCGGGGCCACTCCAAAGAGAAAGATCTCTGGGCTCAGGAGGGGAC 713

RESULT 12
 CD351162
 LOCUS
 DEFINITION UI-M-FY0-cft-p-13-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
 IMAGE: 6852446 5', mRNA sequence.

ACCESSION CD351162
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Mus musculus (house mouse)

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 JOURNAL 1 (bases 1 to 699)
 COMMENT NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mouse1.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1. .699
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6852446"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP FYO"
/notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGGACAGAC. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

BASE COUNT 135 a 221 c 202 g 139 t 2 others

ORIGIN

Query Match 11.1% Score 685.8; DB 14; Length 699;
Best Local Similarity 98.7%; Pred. No. 4.9e-105;
Matches 690; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2753 ATTCTTCAAGATGGGAGCCGACACAGAGTGGGGCTGGGCTTTTCCCGCTTGCGAGG 2812
DB 1 ATTCTTCAAGATGGGAGCCGACACAGAGTGGGGCTGGGCTTTTCCCGCTTGCGAGG 60
QY 2813 ATGTGATTACTCTGCAAGAGACATTTCTCAGACTCTTTGGATGTGCACGCCGTCAAGGCA 2872
DB 61 ATGTGATTACTCTGCAAGAGACATTTCTCAGACTCTTTGGATGTGCACGCCGTCAAGGCA 120
QY 2873 TCATCGTGGGTGTCTGCTGGGCTTTCTCTGCTCTGCTGGGCTGCATGTGTGGCTTAC 2932
DB 121 TCATCGTGGGTGTCTGCTGGGCTTTCTCTGCTCTGCTGGGCTGCATGTGTGGCTTAC 180
QY 2933 GACAAAGCTCCACAGAGGAGCCCTTCCCGATTTGCTCTCAGGACCCCGAGGAACC 2992
DB 181 GACAAAGCTCCACAGAGGAGCCCTTCCCGATTTGCTCTCAGGACCCCGAGGAACC 240
QY 2993 CAGCGCTTACAAAGAGCTCGGCTTGGGCTCCAGTGTCCCTGTGCTCCCATGAGTTGG 3052
DB 241 CAGCGCTTACAAAGAGCTCGAATTGGGCCCCCAGTGTCCCTGTGCTCCCATGAGTTGG 300
QY 3053 AGTCCCTCGTGCATCTCTGCTCCCGATTTGGTCCCCCAGGATTCAGATGTGGAGACA 3112
DB 301 AGTCCCTCGTGCATCTCTGCTCCCGATTTGGTCCCCCAGGATTCAGATGTGGAGACA 360
QY 3113 AGGTGAAGTACACAGACTTATGGGTGGCAGTGTTCAGATTTGCGGGGGCCACTCCAAAGA 3172
DB 361 AGGTGAAGTACACAGACTTATGGGTGGCAGTGTTCAGATTTGCGGGGGCCACTCCAAAGA 420
QY 3173 GAAAGATCTCTCGGCTCAGGAGGGGACCAACTCGGCGAGGCTCTCGGCGAGGCTGTG 3232
DB 421 GAAAGATCTCTCGGCTCAGGAGGGGACCAACTCGGCGAGGCTCTCGGCGAGGCTGTG 480

QY 3233 AGCTGCCCGAGGTAGTGTCAAGCGCGGCTCTGACCGGCTCTGCTGCTCCAGCGG 3292
DB 481 AGCTGCCCGAGGTAGTGTCAAGCGCGGCTCTGACCGGCTCTGCTGCTCCAGCGG 540
QY 3293 GAACCGGGCAGACACTGCTGCTGCAAGCCCTGCTGTATGACGCGCATTAAGACGACGGA 3352
DB 541 GAACCGGGCAGACACTGCTGCTGCAAGCCCTGCTGTATGACGCGCATTAAGACGACGGA 600
QY 3353 GAAAGAGCGGTCGCCAGCGCTGCAGGAATCAGGTGGAAGCTGAGGTGATTTGCCACTCCG 3412
DB 601 GAAAGAGCGGTCGCCAGCGCTGCAGGAATCAGGTGGAAGCTGAGGTGATTTGCCACTCCG 660
QY 3413 ACTTGGGTGCATCCAAAGGATGCTCTGACCTCCACTCC 3451
DB 661 ACTTGGGTGCATCCAAAGGATGCTCTGACCTCCACTCC 699

RESULT 13

BI185886

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

BI185886 900 bp mRNA linear EST 10-JUL-2001
UNL-P-FN-df-h-02-0-UNL-al UNL-P-FN Sus scrofa cDNA clone
UNL-P-FN-df-h-02-0-UNL 3', mRNA sequence.
BI185886
BI185886.1 GI:14660295
EST.
Sus scrofa (pig)
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
Caetano, A.R., Johnson, R.K. and Pomp, D.
Generation and sequence characterization of a normalized cDNA
library from swine ovarian follicles
Mamm. Genome 14 (1), 65-70 (2003)
22419904
12532269
Contact: Pomp, D
Department of Animal Science
University of Nebraska, Lincoln
Lincoln, NE 68583-0908, USA
Tel: 402 472 6416
Fax: 402 472 6362
Email: dpomp@unl.edu
Oligo-dT track not found, Not I site shown in beginning of sequence
is likely internal to the message.
Seq primer: M13 -29
POLYA=No.

FEATURES

Source

1. .900
/organism="Sus scrofa"
/mol_type="mRNA"
/strain="University of Nebraska, Lincoln Swine Selection
Lines"
/db_xref="taxon:9823"
/clone="UNL-P-FN-df-h-02-0-UNL"
/dev_stage="ADULT"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UNL-P-FN"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UNL-P-FN
library is a normalized library representing porcine
ovarian follicles, ranging between 2.0 to 10.0 mm in
diameter, collected during 7 days of the follicular phase
of the pig estrous cycle. This library was derived from
the library UNL-P-F2. The tag is a string of 5-6
nucleotides present between the Not I site and the
oligo-dT track. The library was constructed as described
by Bonaldo, Lennon and Soares, Genome Research 6: 791-806
, 1996.

BASE COUNT 149 a 304 c 291 g 150 t 6 others

ORIGIN

Query Match 10.4%; Score 639.6; DB 12; Length 900;
Best Local Similarity 86.1%; Pred. No. 2.7e-97;
Matches 719; Conservative 0; Mismatches 115; Indels 1; Gaps 1;

QY 739 GCAGCCCGAGAGACACACAGGTAGTGTCTGACAGATGTAGTATGAGTGTGCTGCC 798
Db 7 GCTGCCCCGGAGAACACACAGTGTGTGCGGCCAGAGCGTGTATGAGTGTGCTGCC 66
QY 799 TCTGTGACCCACACCCCTTTTGTCTCTGGTCCGACAGGATGGAAGCCTATCTCCACG 858
Db 67 TCGGCTGACCCACACCCCTATCTGTCTGGTCCGACAGGAGCGGAGCCTATTCTACC 126
QY 859 GATGTATCTGTTCTGGCGGAGCAATCTATCTATCTGTCGACGCGCAGCCTGCGACTCT 918
Db 127 GATGTATCTGTTCTGGCGGAGCAATCTATCTATCTGTCATCACCAGCGCCANCC-CGGACTCT 185
QY 919 GGAGTCTATGTCGCGAGCCCAACAGCCCTCAGCGTGTATCTGCCACTGCGGCTCT 978
Db 186 GCGGTCTAGTCTGACAGAGCAACAGCCCTGATCGCGACTTTTGCCACCGCCGCC 245
QY 979 GAGCTCCGAGTCTTGTCTGCCCCAGCCATCTCGCAGGACCCGAGCGCTCTCGCGGAGC 1038
Db 246 GAGCTCCGCGTCTGCGGCTCTGCGCATCTCGAGGCGCCGAGCGCTATCTCGAGC 305
QY 1039 CGGCGCAGCACCGCGCTCTGTGTGCGGGCGTCCGGGAGCCACGCGCCGCTGCGAC 1098
Db 306 AGGGCGAGCACCGCGCTCTGTGTGCGCGCAACCGGTGAGCGCGCGCGCTGCGG 365
QY 1099 TGGCTGACAGCGGATCCGTTGCGACCAATGGCGGCTCAAGTGCAGGCGCGTGC 1158
Db 366 TGGCTGCAACCGGGCGCGCTGCGGCGCAACCGGGCGTCAAGTGCAGGCGCGCGC 425
QY 1159 GCGAGTCTGCTCATCACTCAGATCGGCTGACGAGCGTGGCTACTACAGTGGGTAGCA 1218
Db 426 AGCAGCTTGTCTACGAGATCGGCTGACAGATGCGGCTACTACAGTGGGTGCGC 485
QY 1219 GAAACAGCGGGAACTGCTGTGCGCTGCGCCCTGGCGGTAGTGTGCGGAGGG 1278
Db 486 GAGAACAGCGGGCACCGGTGCGCGCGCGCTGGCAGTAGTAGTGTGCGGAGGT 545
QY 1279 CTGCCCCAGGCGCCGCTGCGGTACAGCAGCGCGCTGAGCAGTCTCTGTGTGTTG 1338
Db 546 CTGCCCCAGGCGCCGCGGTACCGCGGTACCGCCAGCGCCCTGAGCAGTCTTGTGTG 605
QY 1339 GCCTGGAGCGCTGAGTTGACAGCGAGCAATCATTTGGCTCTCTCTTCACTACCAA 1398
Db 606 GCCTGGAGCGCGCGAGTGTGACAGCGAGCATCTCGGTCTCTCTTCACTACCAA 665
QY 1399 AAGCAAGGGAGTGGCAATGTGGAGTACCAGTTTGCAGTAAACAATGACACACAGAG 1458
Db 666 AAGCAAGGGAGTGGCAATGTGGAGTACCAGTTTGCAGTAAACAATGACACACAGAG 725
QY 1459 CTGAGGTTTGGGACCTGGAACCCCAACCGGATATGATGTTCTAGTGTGCGCTACTCC 1518
Db 726 CTACAGGTTTGGGACCTGAGGCCCAACACAGACTATGATTTCTATGTGTGCGCTACTCA 785
QY 1519 CAGCTGGGGCGCAGCGGACCTCCAGCCCGCTGTGTGCTATACACTGAGCGATG 1573
Db 786 CAGCTGGGGCGCAGCGGACCTCCACCCCGCGCTGGTCCACACATGGATGATG 840

RESULT 14
CB518504
LOCUS CB518504 631 bp mRNA linear EST 28-MAR-2003
DEFINITION UI-M-GH0-ccc-h-04-0-UI-r1 NIH_BMAP_GH0 Mus musculus cDNA clone
IMAGE: 6837653 5', mRNA sequence.
ACCESSION CB518504
VERSION CB518504.1 GI:29351859
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 631)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mouseefi.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

FEATURES
source

Location/Qualifiers

1..631
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6837653"
/tissue_type="Whole brain"
/dev_stage="1, 5, and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP GH0"

Site 1: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAATCAAT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT
ORIGIN

151 a 183 c 180 g 117 t

Query Match
Best Local Similarity
Matches 630; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

10.2%; Score 629.4; DB 14; Length 631;
99.8%; Pred. No. 1.5e-95;

QY 1406 GGGGAGTGGACAATGTGGAGTACCAGTTTGCAGTAAACAATGACACACAGAGCTGCAGG 1465
Db 1 GGGGAGTGGACAATGTGGAGTACCAGTTTGCAGTAAACAATGACACACAGAGCTGCAGG 60

QY 1466 TTCGGGACCTTGAACCCCAACAGGATATGATGTTCTAGTGTGGCTTACTCCAGCTGG 1525
Db 61 TTCGGGACCTTGAACCCCAACAGGATATGATGTTCTAGTGTGGCTTACTCCAGCTGG 120

QY 1526 GGGGACCGCGAACCTCCAGCCCGCTGGTGCATACACTGACGATGTCCCGCGCAG 1585
Db 121 GGGGACCGCGAACCTCCAGCCCGCTGGTGCATACACTGACGATGTCCCGCGCAG 180

QY 1586 CACCCAGCTTACCTTGTTCAGCCCAACCCCTCGGACATCAGGTTGCGATGGCTGCC 1645
Db 181 CACCCAGCTTACCTTGTTCAGCCCAACCCCTCGGACATCAGGTTGCGATGGCTGCC 240

QY 1646 TGCCCTCCAGCTGAGCAATGACAGGTGCTGAAAGTACAAGATAGAGTACGCTTTGGGGA 1705
Db 241 TGCCCTCCAGCTGAGCAATGACAGGTGCTGAAAGTACAAGATAGAGTACGCTTTGGGGA 300

QY 1706 AGGAGATCAGGTTTCTCCACCGGCTGCTCGAAATGACACACAACTTACGTTAACT 1765
Db 301 AGGAGATCAGGTTTCTCCACCGGCTGCTCGAAATGACACACAACTTACGTTAACT 360

QY 1766 CACTTCAGCCAAACAAAGTGTACCGAGTCCGGATTTCAGCTGCGCAGCTGGCGCTG 1825

[illegible]

RESULT 15	BY719391	671 bp	linear	EST 17-DEC-2002
LOCUS	BY719391	Riken full-length enriched, 12 days embryo male wolffian		
DEFINITION	duct Mus musculus cDNA Clone 672045A10 5', mRNA sequence.			
ACCESSION	BY719391			
VERSION	BY719391.1	GI:27132508		
KEYWORDS	EST.			
SOURCE	Mus musculus	(house mouse)		
ORGANISM	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;			

```
QY 5413 ACACATAAGATTATTAATGCTGGCTTCGTATCTCATTAAGACAGGATGTCACTTGAAC 5472
Db |||||
QY 181 ACACATAAGATTATTAATGCTGGCTTCGTATCTCATTAAGACAGGATGTCACTTGAAC 240
Db |||||
QY 5473 TACTTCTATAGATTCAAGTGGCCACGGCCCAACCAACCGTATGTTTCTTCATTGCTCT 5532
Db |||||
QY 241 TACTTCTATAGATTCAAGTGGCCACGGCCCAACCAACCGTATGTTTCTTCATTGCTCT 300
Db |||||
QY 5533 GAAGGTCAAGAGCCTCATTTTGTCTTCTGGTTAGATTCTTCTCCTTGCCTTGAATG 5592
Db |||||
QY 301 GAAGGTCAAGAGCCTCATTTTGTCTTCTGGTTAGATTCTTCTCCTTGCCTTGAATG 360
Db |||||
QY 5593 AAATAACCGTTTTAAACAGTAGGCTCTTAGCATCACACACATAGTCAATTCCTCATGTTCT 5652
Db |||||
QY 361 AAATAACCGTTTTAAACAGTAGGCTCTTAGCATCACACACATAGTCAATTCCTCATGTTCT 420
Db |||||
QY 5653 TGTTTAAACAGCATTTGAGGTTCTGGGTTTAAATTAATAGCTGCAAAATGAGACAATTTA 5712
Db |||||
QY 421 TGTTTAAACAGCACT--GAGGTTCTGGTTTAAATTAATAGCTGCAAAATGAGACAATTTA 478
Db |||||
QY 5713 TAACCCATTAGCTGGTGGAAATTTG--TTCTCAAAAGCAAAATAGTAATAAATCTGGTA 5771
Db |||||
QY 479 TAACCCATTAGCTGGTGGAAATTTGTTCTCAAAAGCAAAATAGTAATAAATCTGGTA 538
Db |||||
QY 5772 TCTGCTATAACTCACAGTTGATAGAAAAGTAGCCAGAACTCACATAGCAATTATATGAT 5831
Db |||||
QY 539 TCTGCTATAACTCACAGTTGATAGAAAATAGCCAGAACTCACATAGCAATTATATGAT 598
Db |||||
QY 5832 TGGGGTTCAGAGTAACTGGGAGTGTAGCTTTGTGACTTTGTAGCACCGGCTTTATTA 5891
Db |||||
QY 599 TTGGGTTCTGAGTAACTGGGAGTGTAGCTTTGTGAAATTTGGAACAACAGGCTCTTATTA 658
Db |||||
QY 5892 GGAAGTCTGTTG 5904
Db |||||
QY 659 AGAAAAGTTGGTG 671
```

Search completed: October 15, 2003, 11:38:58
Job time : 14070.8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 06:22:55 ; Search time 1163.58 Seconds
(without alignments)
14327.989 Million cell updates/sec

Title: US-09-754-997A-1
Perfect score: 6176
Sequence: 1 atggcgggcggaacggg.....taataaactatttccacagac 6176

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_13Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6176	100.0	6176	22	Mouse Nope (neighb
2	2804.2	45.4	3753	22	Nucleotide sequenc
3	2796	45.3	2796	22	Mouse Nope (neighb
4	2681.8	43.4	3741	24	DNA encoding human
5	2118.8	34.3	3450	24	Human polynucleoti
6	1522.4	24.7	2223	22	Nucleotide sequenc
7	825	13.4	825	22	Mouse Nope (neighb
8	337.2	5.5	680	21	Human secreted pro

9	316.6	5.1	2976	25	ABX71437	Human cDNA encodin
10	316.6	5.1	3210	25	ABX71432	Human cDNA encodin
11	316.6	5.1	3219	25	ABX71440	Human cDNA encodin
12	316.6	5.1	3874	25	ABX71447	Human cDNA encodin
13	316.4	5.1	2700	24	AAD36777	Human immunoglobul
14	315.4	5.1	3453	25	ABX71435	Human cDNA encodin
15	289	4.7	2695	23	AAS76339	DNA encoding novel
16	288	4.7	288	22	AAD10030	Mouse Nope (neighb
17	252	4.1	252	22	AAD10032	Mouse Nope (neighb
18	249	4.0	249	22	AAD10028	Mouse Nope (neighb
19	249	4.0	249	22	AAD10029	Mouse Nope (neighb
20	246	4.0	246	22	AAD10031	Mouse Nope (neighb
21	243	3.9	243	22	AAD10024	Mouse Nope (neighb
22	217	3.5	2481	25	ABX71439	Human cDNA encodin
23	217	3.5	2715	25	ABX71434	Human cDNA encodin
24	217	3.5	2724	25	ABX71441	Human cDNA encodin
25	217	3.5	2958	25	ABX71436	Human cDNA encodin
26	213.8	3.5	2139	25	ABX71442	Human cDNA encodin
27	213.8	3.5	2382	25	ABX71445	Human cDNA encodin
28	201.8	3.3	563	24	ABQ25578	Oligonucleotide fo
29	201.8	3.3	563	24	ABQ25579	Oligonucleotide fo
30	200.2	3.2	3330	23	AAS76658	DNA encoding novel
31	195	3.2	195	22	AAD10027	Mouse Nope (neighb
32	192	3.1	192	22	AAD10025	Mouse Nope (neighb
33	190.4	3.1	563	24	ABQ25580	Oligonucleotide fo
34	190.4	3.1	563	24	ABQ25581	Oligonucleotide fo
35	189	3.1	189	22	AAD10026	Mouse Nope (neighb
36	151.2	2.4	1875	25	ABX71443	Human cDNA encodin
37	119	1.9	909	25	ABX71438	Human cDNA encodin
38	119	1.9	1143	25	ABX71433	Human cDNA encodin
39	117	1.9	1644	25	ABX71444	Human cDNA encodin
40	117	1.9	1887	25	ABX71446	Human cDNA encodin
41	99.4	1.6	2450	23	AAS76657	DNA encoding novel
42	93.8	1.5	2560	25	ABX71188	Novel human cDNA s
43	85.6	1.4	439	24	ABS69793	Novel murine polyn
44	75	1.2	370	22	AAS33787	Human cDNA encodin
45	71	1.1	4277	21	AAA37039	Human PRO1190 (UNQ

ALIGNMENTS

RESULT 1

AAD10021
ID AAD10021 standard; cDNA; 6176 BP.
XX
AC AAD10021;
XX
DT 12-SEP-2001 (first entry)
XX Mouse Nope (neighbour of punc ell) cDNA.

XX Mouse; Nope; neighbour of punc ell; cytostatic; neuroprotective; vaccine;
XX gene therapy; cerebroprotective; colonic cancer; mental retardation;
XX tumour suppressor; chromosome 9; transgenic animal; genetic disorder;
XX obesity; Bardet-Biedl syndrome; autosomal recessive disorder; retinitis;
XX polydactyly; pigmentosa; hypogonadism; hypertension; diabetes mellitus;
XX renal anomaly; cardiovascular anomaly; ss.

OS Mus musculus.

Key	Location/Qualifiers
FT CDS	1..3759
FT	/tag= a
FT	/product= "Mouse Nope protein"
FT	1..63
FT sig_peptide	/tag= b
FT	64..3756
FT mat_peptide	/tag= c
FT	/product= "Mature mouse Nope protein"
FT misc_feature	1..3756
FT	/tag= d
FT	/note= "This region is specifically referred as

FT SEQ ID NO: 45 in the specification"
 FT 64..2859
 FT /tag= e
 FT /note= "Extracellular domain"
 FT misc_feature
 FT 139..381
 FT /tag= f
 FT /label= Ig1
 FT /note= "Immunoglobulin domain 1"
 FT misc_feature
 FT 463..654
 FT /tag= g
 FT /label= Ig2
 FT /note= "Immunoglobulin domain 2"
 FT misc_feature
 FT 766..954
 FT /tag= h
 FT /label= Ig3
 FT /note= "Immunoglobulin domain 3"
 FT misc_feature
 FT 1039..1233
 FT /tag= i
 FT /label= Ig4
 FT /note= "Immunoglobulin domain 4"
 FT misc_feature
 FT 1285..1533
 FT /tag= j
 FT /label= Fn1
 FT /note= "Fibronectin-type III domain 1"
 FT misc_feature
 FT 1579..1827
 FT /tag= k
 FT /label= Fn2
 FT /note= "Fibronectin-type III domain 2"
 FT misc_feature
 FT 1888..2175
 FT /tag= l
 FT /label= Fn3
 FT /note= "Fibronectin-type III domain 3"
 FT misc_feature
 FT 2248..2493
 FT /tag= m
 FT /label= Fn4
 FT /note= "Fibronectin-type III domain 4"
 FT misc_feature
 FT 2542..2793
 FT /tag= n
 FT /label= Fn5
 FT /note= "Fibronectin-type III domain 5"
 FT misc_feature
 FT 2932..3756
 FT /tag= o
 FT /note= "Intracellular domain"
 FT 3'UTR
 FT 3819..5682
 FT /tag= p
 FT WO200149714-A2.
 FT
 PD 12-JUL-2001.
 XX
 PD 26-OCT-2000; 2000WO-US29698.
 XX
 PR 04-JAN-2000; 2000US-0174496.
 PR 19-MAY-2000; 2000US-0205789.
 XX
 PA (NEUR-) NEUROSCIENCES RES FOUND INC.
 XX
 PI Salbaum JM;
 XX
 DR WPI; 2001-441846/47.
 DR P-PSDB; AAE05251.
 XX
 PT Murine Nope polypeptides and nucleic acids useful for preventing,
 PT diagnosing and treating colonic cancer and Bardet-Biedl syndrome -
 XX Claim 10; Fig 2A; 99pp; English.
 XX
 PS The present invention relates to Nope (neighbour of punc ell) which is
 CC used in the prevention, treatment and diagnosis of diseases associated
 CC w/ inappropriate Nope expression such as cancers especially colonic
 CC cancer and genetic disorders, as Nope is thought to be a tumour
 CC suppressor. Nope gene is located on chromosome 9 and is used in gene
 CC therapy. Nope is used as vaccine. Nope gene may be administered to treat

CC diseases by rectifying mutations or deletions in a patient's genome that
 CC affect the activity of Nope by expressing inactive proteins or to
 CC supplement the patients own production of Nope polypeptides. Nope gene
 CC is used to study the expression and function of Nope polypeptides and
 CC their role in metabolism through the creation of transgenic animal
 CC models. The anti- Nope antibodies and Nope antagonists may also be used
 CC to down regulate Nope expression and activity for the treatment of
 CC Bardet-Biedl syndrome which is an autosomal recessive disorder
 CC characterised by mental retardation, obesity, polydactyly, retinitis
 CC pigmentosa and hypogonadism. Patients with Bardet-Biedl syndrome have a
 CC high incidence of hypertension, diabetes mellitus and renal and
 CC cardiovascular anomalies. The present sequence is mouse Nope (neighbour
 CC of punc ell) cDNA.
 XX

XX Sequence 6176 BP; 1316 A; 1756 C; 1634 G; 1470 T; 0 other;

Query Match 100.0%; Score 6176; DB 22; Length 6176;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 6176; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGCGCGGGCGGACACGGGGCGGGCTCTGGTGTGACCTTCTGCTGCTGTCCGGG 60
 DB |||||
 QY 1 ATGGCGCGGGCGGACACGGGGCGGGCTCTGGTGTGACCTTCTGCTGCTGTCCGGG 60
 DB |||||
 QY 61 CGCGGGAGCTGCCATTGCCCGGAGGACAACTGTCAAGCTGAGCTGTGATGAGGACCC 120
 DB |||||
 QY 61 CGCGGGAGCTGCCATTGCCCGGAGGACAACTGTCAAGCTGAGCTGTGATGAGGACCC 120
 DB |||||
 QY 121 CTGCAAGTATCTCTGGGCTCTGAGCAGGCTGTGGTGTGAGTGTGACCTTTGGGGGCTACA 180
 DB |||||
 QY 121 CTGCAAGTATCTCTGGGCTCTGAGCAGGCTGTGGTGTGAGTGTGACCTTTGGGGGCTACA 180
 DB |||||
 QY 181 GCTGTGGGCTCTCCGACAGGCTGACATGGACCAAGGATGGAGACACTGTACTAGAGCAT 240
 DB |||||
 QY 181 GCTGTGGGCTCTCCGACAGGCTGACATGGACCAAGGATGGAGACACTGTACTAGAGCAT 240
 DB |||||
 QY 241 GAGAACCTGTCACCTGCTACCCCAATGGCTCCCTGTGGTGTCTCTCAACCTTAGAGCAAGAA 300
 DB |||||
 QY 241 GAGAACCTGTCACCTGCTACCCCAATGGCTCCCTGTGGTGTCTCTCAACCTTAGAGCAAGAA 300
 DB |||||
 QY 301 GACAGGATGATGAGGAGCTCTTAGGATCTGGAGGTCATCTAGAGGAGCTATTCCTGT 360
 DB |||||
 QY 301 GACAGGATGATGAGGAGCTCTTAGGATCTGGAGGTCATCTAGAGGAGCTATTCCTGT 360
 DB |||||
 QY 361 CTGGGCCACAGCCCGCTAGGAGTGTGGCCAGCAGTTGTGTCTCAAGCTTGCACCA 420
 DB |||||
 QY 361 CTGGGCCACAGCCCGCTAGGAGTGTGGCCAGCAGTTGTGTCTCAAGCTTGCACCA 420
 DB |||||
 QY 421 CTCGAAGACTTCTCTTGACCCCGAGTCCAGATTTGTGGAGGAGAAACGGGACAGCACGC 480
 DB |||||
 QY 421 CTCGAAGACTTCTCTTGACCCCGAGTCCAGATTTGTGGAGGAGAAACGGGACAGCACGC 480
 DB |||||
 QY 481 TTTGAAATGCCACACCAAGGGCTTTCAGCCCCCATCATTTACTTGGGAAAGAGCACAGGTG 540
 DB |||||
 QY 481 TTTGAAATGCCACACCAAGGGCTTTCAGCCCCCATCATTTACTTGGGAAAGAGCACAGGTG 540
 DB |||||
 QY 541 ACCGTGCTGAGGAGCCCGGCTCATCTCTTCCCAAGTGGCTCTCCAGATCTCTAGAT 600
 DB |||||
 QY 541 ACCGTGCTGAGGAGCCCGGCTCATCTCTTCCCAAGTGGCTCTCCAGATCTCTAGAT 600
 DB |||||
 QY 601 GTCCAGGACAGTGTGATGCGAGGCTCTTACCGCTCGGTGGCCCAACAAATTCAGCCCGCCACCA 660
 DB |||||
 QY 601 GTCCAGGACAGTGTGATGCGAGGCTCTTACCGCTCGGTGGCCCAACAAATTCAGCCCGCCACCA 660
 DB |||||
 QY 661 TTTAGCCAGAGGCTCTGCTCATCTGTGGCCCTCAGAGGCTCTTTGAGGCTTACCAGGGGG 720
 DB |||||
 QY 661 TTTAGCCAGAGGCTCTGCTCATCTGTGGCCCTCAGAGGCTCTTTGAGGCTTACCAGGGGG 720
 DB |||||
 QY 721 CAGGATGTGCTCATTTGTGGCAGCCCGGAGAGACCAACGCTAGTCTCTGGACAGATGTA 780
 DB |||||
 QY 721 CAGGATGTGCTCATTTGTGGCAGCCCGGAGAGACCAACGCTAGTCTCTGGACAGATGTA 780
 DB |||||
 QY 781 GTGATGGAGTGGCTGGCCCTCTGCTGACCCACCCCTTTTGTGTCTCTGGGTCGACAGGAT 840

Db 781 GTGATGGAGTGGTGGCTCTGCTGACCCACCCCTTTTGTGCTCGGTCCGACGAT 840
QY 841 GGAAGCCTAICTCAACGATGTCACTGTTCTGGCCGAGACCAATCTACTCATGCCAGC 900
Db 841 GGAAGCCTAICTCAACGATGTCACTGTTCTGGCCGAGACCAATCTACTCATGCCAGC 900
QY 901 GCGAGCCTCGGCATCTGAGTCTATGTCTGCCGAGCAACAAGCCCTCATCGGTGAC 960
Db 901 GCGAGCCTCGGCATCTGAGTCTATGTCTGCCGAGCAACAAGCCCTCATCGGTGAC 960
QY 961 TTCCGCACTGGCGCTGCTGAGCTCCGAGTGTCTGCTGCCAGCCATCTCCGAGSCACC 1020
Db 961 TTCCGCACTGGCGCTGCTGAGCTCCGAGTGTCTGCTGCCAGCCATCTCCGAGSCACC 1020
QY 1021 GAGGCGCTCTCGCGACCGCGGCCAGACCGCGCTTCTGTTGCCGGCGTCCGGGGAG 1080
Db 1021 GAGGCGCTCTCGCGACCGCGGCCAGACCGCGCTTCTGTTGCCGGCGTCCGGGGAG 1080
QY 1081 CCAAGGCGCGCTGCTGAGTGTGCAAGCGGGATCCCGTTGCGACCCCAATGGCGGCTC 1140
Db 1081 CCAAGGCGCGCTGCTGAGTGTGCAAGCGGGATCCCGTTGCGACCCCAATGGCGGCTC 1140
QY 1141 AAGGTGCAAGGCGGCTGCGGAGCTTGTCTCATCTCAGATCGGCTCGAGGACGCTGGC 1200
Db 1141 AAGGTGCAAGGCGGCTGCGGAGCTTGTCTCATCTCAGATCGGCTCGAGGACGCTGGC 1200
QY 1201 TACTACGAGTGTAGCAGAAAAACAGCGCGGAACTGCTGTGCGCTGCGCCCCGCGC 1260
Db 1201 TACTACGAGTGTAGCAGAAAAACAGCGCGGAACTGCTGTGCGCTGCGCCCCGCGC 1260
QY 1261 GTAGTGTGCGGAGGCGTGCAGCGCCGCACTCGGGTCAAGCCACCGCGCTGAGC 1320
Db 1261 GTAGTGTGCGGAGGCGTGCAGCGCCGCACTCGGGTCAAGCCACCGCGCTGAGC 1320
QY 1321 AGCTCTCTGTGCTGCTGGAGCGGCTGAGTTGCAAGCGAGCAATCAATTGGC 1380
Db 1321 AGCTCTCTGTGCTGCTGGAGCGGCTGAGTTGCAAGCGAGCAATCAATTGGC 1380
QY 1381 TTCTCTCTTCACTACAAAGGCAAGGGAGTGGACAAATGTGGAGTACAGTTTGCAGTA 1440
Db 1381 TTCTCTCTTCACTACAAAGGCAAGGGAGTGGACAAATGTGGAGTACAGTTTGCAGTA 1440
QY 1441 AACAAATGACACACAGAGCTGAGGTTCCGGACCTTGGAACTTGGAACTTGGAGTTG 1500
Db 1441 AACAAATGACACACAGAGCTGAGGTTCCGGACCTTGGAACTTGGAACTTGGAGTTG 1500
QY 1501 TACGTGTGGCTACTCCAGCTGGGGGCGAGCCGAACTCCAGCCAGCCCTGGTGCAT 1560
Db 1501 TACGTGTGGCTACTCCAGCTGGGGGCGAGCCGAACTCCAGCCAGCCCTGGTGCAT 1560
QY 1561 ACACGTGACGATGTCCCGAGCGACACCCAGCTTACCTTGTTCAGGCCCAACCCCTCG 1620
Db 1561 ACACGTGACGATGTCCCGAGCGACACCCAGCTTACCTTGTTCAGGCCCAACCCCTCG 1620
QY 1621 GACATCAGGGTGGCATGCTGCCCTGCCCTCCAGCTTGGAACTTGGAGTGTGAG 1680
Db 1621 GACATCAGGGTGGCATGCTGCCCTGCCCTCCAGCTTGGAACTTGGAGTGTGAG 1680
QY 1681 TACAAGATAGAGTACGGTTTGGGAGGAGAAATCAGGTTTCTCCACCGAGGTGCTGGA 1740
Db 1681 TACAAGATAGAGTACGGTTTGGGAGGAGAAATCAGGTTTCTCCACCGAGGTGCTGGA 1740
QY 1741 AATGAGACAACTTACGTTAACTCAGTTCAAGCCAAACAAAGTGTACCGAGTCCGAT 1800
Db 1741 AATGAGACAACTTACGTTAACTCAGTTCAAGCCAAACAAAGTGTACCGAGTCCGAT 1800
QY 1801 TCAGTGGCACTGGCGCTGGCTATGAGTCCCTTCTCAGTGGATGACGACACACCT 1860
Db 1801 TCAGTGGCACTGGCGCTGGCTATGAGTCCCTTCTCAGTGGATGACGACACACCT 1860
QY 1861 GGTGTGCAACACGAGCCATGTTCCCTTGGCCCTGCAGAAATGGAAGGTGAGGCAAG 1920

Db 1861 GGTGTGCAACACGAGCCATGTTCCCTTGGCCCTGCAGAAATGGAAGGTGAGGCAAG 1920
QY 1921 ATGGAGTCCCTGTGTGTATGCGACCGCCCCCTTCACCCCACCCAGATCTCTGGATAC 1980
Db 1921 ATGGAGTCCCTGTGTGTATGCGACCGCCCCCTTCACCCCACCCAGATCTCTGGATAC 1980
QY 1981 AAATCTTACTTGGGGAGAGTGGGAAACAGAGGAGGAGGAGATGTGTGACCGCCCCCAGG 2040
Db 1981 AAATCTTACTTGGGGAGAGTGGGAAACAGAGGAGGAGGAGATGTGTGACCGCCCCCAGG 2040
QY 2041 GGTCTGAGATCAAGCTTGGGACGTGGGGCCGTGCGGCTGGAAGAAAGTGAAGCAG 2100
Db 2041 GGTCTGAGATCAAGCTTGGGACGTGGGGCCGTGCGGCTGGAAGAAAGTGAAGCAG 2100
QY 2101 TATGAACTGACCCAGTGTAGTCCCTGGCAGGCGGTACGAGGTGAAGCTCGTAGCTTCAAC 2160
Db 2101 TATGAACTGACCCAGTGTAGTCCCTGGCAGGCGGTACGAGGTGAAGCTCGTAGCTTCAAC 2160
QY 2161 AAACACAGGACCGCTACGCTGTGTGTGGAAGGGCAAGACGAGAGGCGCCACCGCA 2220
Db 2161 AAACACAGGACCGCTACGCTGTGTGTGGAAGGGCAAGACGAGAGGCGCCACCGCA 2220
QY 2221 GACCTGCTATCCAGAGGCGGCCACCGCTGCTCTGCCCCCATGTCCACGAGCTCAAC 2280
Db 2221 GACCTGCTATCCAGAGGCGGCCACCGCTGCTCTGCCCCCATGTCCACGAGCTCAAC 2280
QY 2281 AGCTTCACTTCCATTTGGCTTGGTGAAGAACGACGACTTTACCACTGTCAAGATGTCTC 2340
Db 2281 AGCTTCACTTCCATTTGGCTTGGTGAAGAACGACGACTTTACCACTGTCAAGATGTCTC 2340
QY 2341 AACTACTGTATCGCTTGGGCGCTCAGGAAATGTTCCCTGGTCACTACTAT 2400
Db 2341 AACTACTGTATCGCTTGGGCGCTCAGGAAATGTTCCCTGGTCACTACTAT 2400
QY 2401 ACCAGCTCTGAGAAAGACATTTCTCATTTGGCGGCTGAAACATTTACCAAGTACGATTT 2460
Db 2401 ACCAGCTCTGAGAAAGACATTTCTCATTTGGCGGCTGAAACATTTACCAAGTACGATTT 2460
QY 2461 GCGGTACAGTCCCAAGAGTGGATATGGATGGGCGCTTTGGCTCGGTAGAAACGCTCC 2520
Db 2461 GCGGTACAGTCCCAAGAGTGGATATGGATGGGCGCTTTGGCTCGGTAGAAACGCTCC 2520
QY 2521 ACCCTGCTGACCGGCTTCAACACTCTTCTGACCTGTGCGCTGAGCGCCCTGACACCA 2580
Db 2521 ACCCTGCTGACCGGCTTCAACACTCTTCTGACCTGTGCGCTGAGCGCCCTGACACCA 2580
QY 2581 TCCACCGTTCGGTTTACACTGCTGTCGCCCAAGAGGCGCAATGAGTGTGAGATTTGAGAT 2640
Db 2581 TCCACCGTTCGGTTTACACTGCTGTCGCCCAAGAGGCGCAATGAGTGTGAGATTTGAGAT 2640
QY 2641 CTAAATCTTACAGCAACCAACCAAGCGCCGAAACCAAGTGGACACTGCTCAACCA 2700
Db 2641 CTAAATCTTACAGCAACCAACCAAGCGCCGAAACCAAGTGGACACTGCTCAACCA 2700
QY 2701 GAGGAAACATCTTCACTGAGAGTCCATGGCTTACAGAGTACACTCGGTATTTCTTC 2760
Db 2701 GAGGAAACATCTTCACTGAGAGTCCATGGCTTACAGAGTACACTCGGTATTTCTTC 2760
QY 2761 AAGATGGAGCGCGCAAGAGTGGGCGCTGGGCGCTTTTCCGCTTGCAGGATGTGAT 2820
Db 2761 AAGATGGAGCGCGCAAGAGTGGGCGCTGGGCGCTTTTCCGCTTGCAGGATGTGAT 2820
QY 2821 ACTCTGAAAGACATTTCTCAGACTCTTGTGATGTGCAAGCGCGTCAAGGATCATCTGT 2880
Db 2821 ACTCTGAAAGACATTTCTCAGACTCTTGTGATGTGCAAGCGCGTCAAGGATCATCTGT 2880
QY 2881 GGTGTCTGCTGGCGCTTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 2940
Db 2881 GGTGTCTGCTGGCGCTTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 2940
QY 2941 TCCACAGGAGAGGCTTCCCGGATTTGCTCTCTCAGGACCCCGAGGAAACCCAGCGCTC 3000
Db 2941 TCCACAGGAGAGGCTTCCCGGATTTGCTCTCTCAGGACCCCGAGGAAACCCAGCGCTC 3000

Qy	3001	TACAACAGAGCTCGGCTTGGGCCTCCCGAGTGTCCTGTCTGCCCATGATGTTGGAGTCCCTC	3060
Db	3001	TACAACAGAGCTCGGCTTGGGCCTCCCGAGTGTCCTGTCTGCCCATGATGTTGGAGTCCCTC	3060
Qy	3061	GTGCATCCTCGTCCCACAGATTGGTGGCCCCACCACCCCTCAGATGTGGNAGCAAGGCTGAA	3120
Db	3061	GTGCATCCTCGTCCCACAGATTGGTGGCCCCACCACCCCTCAGATGTGGNAGCAAGGCTGAA	3120
Qy	3121	GTAACAAGCCTTATGGGTGGCAGTGTTTTCAGATTGGCCGGGCCACTCAAAGAGAAGAATC	3180
Db	3121	GTACACAGCCTTATGGGTGGCAGTGTTTTCAGATTGGCCGGGCCACTCAAAGAGAAGAATC	3180
Qy	3181	TCCTGGGCTCAGGCACGGGGAGCAAAACTGGGACGGCTTCCTGGGACGGCTGTGAGCTGCCCC	3240
Db	3181	TCCTGGGCTCAGGCACGGGGAGCAAAACTGGGACGGCTTCCTGGGACGGCTGTGAGCTGCCCC	3240
Qy	3241	CAGGGTAGTGGTCCAAGGCGGGCTCTGACCCGTGCTCTGCTGTGCTCCAGCGGGNAACGGG	3300
Db	3241	CAGGGTAGTGGTCCAAGGCGGGCTCTGACCCGTGCTCTGCTGTGCTCCAGCGGGNAACGGG	3300
Qy	3301	CAGACACTGTGCTGAAGCCCTGGTGTATGACGGCATAAAGGACCAACGGGAGAAAGAG	3360
Db	3301	CAGACACTGTGCTGAAGCCCTGGTGTATGACGGCATAAAGGACCAACGGGAGAAAGAG	3360
Qy	3361	CCGTCCCCAGCCTCAGGAATCAGGTTGGAAAGCTGAGTGCATTTGTCCACTCCGACTTCGGT	3420
Db	3361	CCGTCCCCAGCCTCAGGAATCAGGTTGGAAAGCTGAGTGCATTTGTCCACTCCGACTTCGGT	3420
Qy	3421	GCATCCAAAGGATGCTGTGACCTCCACTCCAGACCTCGAGCCAGAGGAACCACTGACT	3480
Db	3421	GCATCCAAAGGATGCTGTGACCTCCACTCCAGACCTCGAGCCAGAGGAACCACTGACT	3480
Qy	3481	GCAGAGACTCTGCCCTCCACGCTCGGAGCTGTGATCTGTCTCAAGGAGCAGACTGGCTG	3540
Db	3481	GCAGAGACTCTGCCCTCCACGCTCGGAGCTGTGATCTGTCTCAAGGAGCAGACTGGCTG	3540
Qy	3541	GGCAGGAGCTGGGAGGGTGCCAACCAACCAACAGTGGGCCAGAGAGGCTCACCTGCTTG	3600
Db	3541	GGCAGGAGCTGGGAGGGTGCCAACCAACCAACAGTGGGCCAGAGAGGCTCACCTGCTTG	3600
Qy	3601	CCAGAAGCAGCAGTGCTCTGTGCTCTGTCTCAGACCTCCAGCCACGATGTGATAGAG	3660
Db	3601	CCAGAAGCAGCAGTGCTCTGTGCTCTGTCTCAGACCTCCAGCCACGATGTGATAGAG	3660
Qy	3661	GAGGCCCTTGGAAAAAGCTGCAGCCCAAAGCCCTGTCTCTAACAGTCAGCCCAAGC	3720
Db	3661	GAGGCCCTTGGAAAAAGCTGCAGCCCAAAGCCCTGTCTCTAACAGTCAGCCCAAGC	3720
Qy	3721	CTTCCCAGGGCCCTGTCTCTCTGTCTCAGGTCCCTTGAGCAGAGGCGATATGGCTCA	3780
Db	3721	CTTCCCAGGGCCCTGTCTCTCTGTCTCAGGTCCCTTGAGCAGAGGCGATATGGCTCA	3780
Qy	3781	GGAAACATGCATGCATGGCTACACATGTGTGTATAGAGATATCCATAAGTCTCTGGAGC	3840
Db	3781	GGAAACATGCATGCATGGCTACACATGTGTGTATAGAGATATCCATAAGTCTCTGGAGC	3840
Qy	3841	CTCTTAGGTCCTTTGGCTGGGGTGGGAGACTTTACTCTCCCTCATATTTCTGGATCA	3900
Db	3841	CTCTTAGGTCCTTTGGCTGGGGTGGGAGACTTTACTCTCCCTCATATTTCTGGATCA	3900
Qy	3901	CATACAGAGGGAATTGAGACACAGCTCTGTGTAAATGGACACGTGTGAAAGTCCGTGTGTGT	3960
Db	3901	CATACAGAGGGAATTGAGACACAGCTCTGTGTAAATGGACACGTGTGAAAGTCCGTGTGTGT	3960
Qy	3961	GTGTGTGTGTGTGTCTGGTGTAGCAAGCTCTGTGTAAATGGACACGTGTGAGCACTCATGTG	4020
Db	3961	GTGTGTGTGTGTGTCTGGTGTAGCAAGCTCTGTGTAAATGGACACGTGTGAGCACTCATGTG	4020
Qy	4021	GCCTAGTTGACCCCTCGGTGGCAGGATGGTGTAAACAGTGATCAGTGGCAGCTCTTTGAGCT	4080
Db	4021	GCCTAGTTGACCCCTCGGTGGCAGGATGGTGTAAACAGTGATCAGTGGCAGCTCTTTGAGCT	4080

[illegible]

Db 5161 GTGGCATCTGCTCTTCTGTCACATCTTCTTATTTCAACTGGCTGGCACATCAAGTGT 5220
Qy 5221 AACTCTGGCTCTGGCCAAAGTTAGAAATAACAGTCTATTTTCCCTTTATTTATTTA 5280
Db 5221 AACTCTGGCTCTGGCCAAAGTTAGAAATAACAGTCTATTTTCCCTTTATTTATTTA 5280
Qy 5281 TTTTATTTTATTTATGCTCTTTCAGTGGAGTTGTAGCTTCTGAAAGCGTCTGTATT 5340
Db 5281 TTTTATTTTATTTATGCTCTTTCAGTGGAGTTGTAGCTTCTGAAAGCGTCTGTATT 5340
Qy 5341 AGCCTGTGTGTCACCTCATGTTTGGACCCCAACCAATTCCTTCTCCCTCTTCAGC 5400
Db 5341 AGCCTGTGTGTCACCTCATGTTTGGACCCCAACCAATTCCTTCTCCCTCTTCAGC 5400
Qy 5401 CAGCCTATGATAAACAATAAGATTAATATGCTGGCTTGTATCTCATTAAGACAGAT 5460
Db 5401 CAGCCTATGATAAACAATAAGATTAATATGCTGGCTTGTATCTCATTAAGACAGAT 5460
Qy 5461 TGTCACTTGAACCTACTTCTATAGCAATTCAGAGTGGCCAGGCAACACACCGTATGTT 5520
Db 5461 TGTCACTTGAACCTACTTCTATAGCAATTCAGAGTGGCCAGGCAACACACCGTATGTT 5520
Qy 5521 CTTCACTGCTGGAAGTCAAGAGCTCATTTTGTGTTTCTGGTGTAGATCTTTTCTCCTC 5580
Db 5521 CTTCACTGCTGGAAGTCAAGAGCTCATTTTGTGTTTCTGGTGTAGATCTTTTCTCCTC 5580
Qy 5581 TTGCTTGAATCAAAATAACCGTTTAAACAGTAGGCTCTTAGCATCACACCATAGTAT 5640
Db 5581 TTGCTTGAATCAAAATAACCGTTTAAACAGTAGGCTCTTAGCATCACACCATAGTAT 5640
Qy 5641 TCCTCATGCTTCTGTTTAAACAGCACTTGAGGCTCTGGGTTTAAATTAAGTGCATA 5700
Db 5641 TCCTCATGCTTCTGTTTAAACAGCACTTGAGGCTCTGGGTTTAAATTAAGTGCATA 5700
Qy 5701 TGAGCAATTTAATACCAATAGGCTGGGTGAAATTTGTTCTCAAAAGCAATTAAGTAA 5760
Db 5701 TGAGCAATTTAATACCAATAGGCTGGGTGAAATTTGTTCTCAAAAGCAATTAAGTAA 5760
Qy 5761 TAAATCTGATCTGCTATTAATCAAGTGTGATTAAGAGTGGAGTGGAGTCTTGTAGTAC 5820
Db 5761 TAAATCTGATCTGCTATTAATCAAGTGTGATTAAGAGTGGAGTGGAGTCTTGTAGTAC 5820
Qy 5821 TTATATGATGATGGGCTCTGAGTAACTGGGAGTGTAGCTTGTGACTTTGTAGTAC 5880
Db 5821 TTATATGATGATGGGCTCTGAGTAACTGGGAGTGTAGCTTGTGACTTTGTAGTAC 5880
Qy 5881 AGGTCTTATAGGAAGTCTGTTGGCTTTTACAGGCAATAGTCCCTTGTGCTGCTG 5940
Db 5881 AGGTCTTATAGGAAGTCTGTTGGCTTTTACAGGCAATAGTCCCTTGTGCTGCTG 5940
Qy 5941 ATGGATGCTTAAAGTCTTGTGGAGTCTCATTTAAGAAATTCCTTTCTCGAGCATGACA 6000
Db 5941 ATGGATGCTTAAAGTCTTGTGGAGTCTCATTTAAGAAATTCCTTTCTCGAGCATGACA 6000
Qy 6001 GTGATGCAATPACTTACATGCTCACTGTTTACCTGGCTAGTTTGTGCTGGGTATT 6060
Db 6001 GTGATGCAATPACTTACATGCTCACTGTTTACCTGGCTAGTTTGTGCTGGGTATT 6060
Qy 6061 AATTGACATTTCCAGCATCATGCTTCTCTTACAAATATGATATTTTATTTGTTACAC 6120
Db 6061 AATTGACATTTCCAGCATCATGCTTCTCTTACAAATATGATATTTTATTTGTTACAC 6120
Qy 6121 TAAGTGTGTATGATGATCTGCTGCTGTTAAGAAATTAATAAATTAATTTTCCAGAC 6176
Db 6121 TAAGTGTGTATGATGATCTGCTGCTGTTAAGAAATTAATAAATTAATTTTCCAGAC 6176

RESULT 2
AAI67204
ID AAI67204 standard; DNA; 3753 BP.
XX
AC AAI67204;

XX 11-FEB-2002 (first entry)
XX Nucleotide sequence of GSK gene Id 27142.
XX Peptide hormone; antidiabetic; anorectic; antianorectic; antiasthmatic;
XX anidipressant; nootropic; neuroprotectant; hypotensive; hypertensive;
XX cytosatic; cerebroprotective; vasotropic; human; ds.
XX Homo sapiens.
XX WO200172961-A2.
XX 04-OCT-2001.
XX 22-MAR-2001; 2001WO-US09226.
XX 24-MAR-2000; 2000US-192158P.
XX 28-MAR-2000; 2000US-192668P.
XX 27-APR-2000; 2000US-200166P.
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX (SMIK) SMITHKLINE BEECHAM PLC.
XX Agarwal P, Murdoch PR, Rizvi SK, Smith RP, Xiang Z, Kabnick KS;
XX Lai Y;
XX WPI; 2001-639223/73.
XX P-P8DB; AAG65914.
XX Isolated polypeptides, which may be peptide hormones, which are
XX identified by high throughput genome-based biology which identifies
XX genes and gene products as therapeutic targets for treatment of
XX diseases such as diabetes and cancer
XX Claim 2; Page 55-56; 99pp; English.
XX The invention provides polypeptides (AAG65986-65918) which may be peptide
XX hormones (including insulin, growth hormones, chemokines, cytokines,
XX neuropeptides, integrins, kallikreins, lamins, melanins, natruiretic
XX hormones, neuropepsin, pituitary hormones, pleiotrophins, prostaglandins,
XX secretogranins, selectins, thromboglobulins, thymosins) identified by
XX high throughput genome-based biology and polynucleotides (AAI67176-67208)
XX encoding them. The polypeptides can be expressed by standard recombinant
XX methodology. The polypeptides are useful in the treatment of disease such
XX as diabetes, breast-, prostate-, colon cancer and other malignant tumors,
XX hyper- and hypotension, obesity, bulimia, anorexia, growth abnormalities,
XX asthma, manic depression, dementia, delirium, mental retardation,
XX Huntington's disease, Tourette's syndrome, schizophrenia, growth, mental
XX or sexual development disorders, and dysfunctions of the blood cascade
XX system including those leading to stroke. The polynucleotides may be used
XX as diagnostic reagents through detecting mutations in the associated gene
XX and for chromosome localization and for tissue expression studies. The
XX polypeptides and polynucleotides may also be used as vaccines.
XX SQ Sequence 3753 BP; 714 A; 1220 C; 1152 G; 667 T; 0 other;
Query Match 45.4%; Score 2804.2; DB 22; Length 3753;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 3171; Conservative 0; Mismatches 568; Indels 6; Gaps 2;
Qy 3 GGCGCGGCGCACACGCGCGGCTCTCTGCTGCTGACCTTCTGCTGCTGCTGCTGCGCG 62
Db 9 GGGGAGCGCGCGCGCGCGGCTCTCTGCTGCTGACCTTCTGCTGCTGCTGCTGCGCG 68
Qy 63 CGGGGAGCTGCTGCTTGGCCCCAGGAGACGACTGTGAGCTGTGATGAGGAGACCCCT 122
Db 69 CGGGGAGCTGCTGCTTGGCCCCAGGAGACGACTGTGAGCTGTGAGTGGGCGCACT 128
Qy 123 GCAAGTGTATCTGGGCGCTGAGAGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 182
Db 129 GCAAGTGTATCTGGGCGCGCGCGGCTGCAAGTGTGCTGTGCTGTGCTGTGCTGTGCTGT 188

QY	183	TGCTGGGCTCCGACCAAGGCTGATCGAGCAAGGATGGAGACATGTCTACTAGACATGA	242		Db	1266	GGTGTGCGGAGGGCTGCCAGCGCCCCACGCGGGTCACTGTCTAGCCACTGACGAG	1325	
Db	189	CGCTGGACCCCCACCAAGGCTGACCTGGAGCAAGGATGGGACACCTCTGTGGAGCACGA	248		QY	1323	CTCCCTCTGTCTGTGGCTGGAGCGGCTGAGTTGCACAGCGAGCAATCATTTGGCTT	1382	
QY	243	GAACCTGCACCTGCTACCAATGGCTCCCTGTGGCTGTCTCACTCCCTAGAGCAAGAGA	302		Db	1326	CTCCCTGTGTGTGGCTGGAGCGGCTCGAGATGCACAGCGAGAGATCATCGGCTT	1385	
Db	249	CCACTTACACCTCTGCCCAATGGTTCCCTGTGGCTGTCCAGGCCACTAGCACCCAAATGG	308		QY	1383	CTCTCTTCACTACCAAAAGGCAAGGGGAGTGGAACAATGTGGAGTACCAAGTTTGCAGTAAA	1442	
QY	303	CAGCGATGATGAGAAAGCTCTTAGGATCTGGAAGTCACTTGAGGGCAGCTATTTCCTGTCT	362		Db	1386	CTCTCTCCACTACCAAGGCAAGGGCACTGGCAATGTGGAAATACCAAGTTTGCAGTGAA	1445	
Db	309	CAG---TGACAGTCTAGTCCCTGAGGCTGTGGGGTCAATTTGAAGGCAACTATTCTGTGCTT	365		QY	1443	CAATGACACCAAGAGCTGCGAGTTTGGGAACTTGGAACCCAAACAGATATGAGTTCTA	1502	
QY	363	GGCCCAAGCCCTAGGAGTGTGGCCAGCCAGGTTGTCTGTGGTCAAGCTTGGCCACACT	422		Db	1446	CAACGACACCAAGAACTACAGTTTGGGACCTGGNAACCAACACAGATATGAGTTCTA	1505	
Db	366	AGCCCAAGCCCTCCGAGTGTGGCAGCCAGACTGTCTGTGGTCAAGCTTCCACACT	425		QY	1503	CGTGTGGCTTACTCCAGCTGGGGGCCAGCGAACCCTCCAGCCAGCGCTGGTGCATAC	1562	
QY	423	CGAAGACTTCTCTGTGACCCCGAGTCCAGATGTGTGGAGGAAACGGGACAGCAAGCTT	482		Db	1506	CGTGTGGCTTACTCCAGCTGGGAGCCAGCGCACCTCCACCCAGCACTGGTGCACAC	1565	
Db	426	CGCAGACTTCTCTGTGACCCCGAGTCTCAGACGGTGGAGGAAACGGGACAGCTGGCTT	485		QY	1563	ACTGACAGATGTCCCGAGCGACACCCAGCTTACTTGTCCAGCCCAACCCCTCGGA	1622	
QY	483	TGAATGCCACCAAGGGCTCCAGCCCCCATATTACTTTGGGAAAGGACCAAGGTGAC	542		Db	1566	ACTGATGATGTCTCCAGTGCAGCACCCAGCTCTCCCTGTCCAGCCCAACCCCTCGGA	1625	
Db	486	TGAGTGCCCAATTGAAGGGCTGCAGTCCCATCATTTACTTTGGGAGAGGACCAAGGTGAC	545		QY	1623	CATCAGGTGGATGGCTGCCCTCCAGCTGCAGCAATGAGCAAGTGTCTGAAAGTA	1682	
QY	543	CGTGCCTGAGGAGCCCGGCTCATCACTCTTCCCAAGTGGCTCCTCCAGATCCTAGATGT	602		Db	1626	CATCAGGGTGGCTGGCTGCCCTGCCCGCCAGCTGCAGCAATGGCGAGTGTGTAAGTA	1685	
Db	546	ATTGCTCTGAGGAGCTCGCTCATCTGCTTCCCAAGCGGCTCTTTCAGATCCTGGATGT	605		QY	1683	CAAGATAGATGAGTGGTGGGGAAGGAGATCAGGTTTCTCCACCGAGGTCCTGGAAA	1742	
QY	603	CAAGGACAGTATGAGAGTCTCTACCGCTGGCTGGGCAACAAATTCAGCCCGCCAAAGATT	662		Db	1686	CAAGATAGATACGGTTTGGGAAAGGAGATCAGATTTTCTCTACTGAGGTGCGAGGAAA	1745	
Db	606	TCAGGAGAGTATGACAGGCGCTTACCGCTGCTGCTCCCAAGCTGCGCCAGCACTT	665		QY	1743	TGAGACACAATCTTACGTTAAACTCACTTCAGGCCAAACAAAGGTGACCGAGTCCGATTTTC	1802	
QY	663	CAGCAGGAGGCTCGCTCACTGTGGCCCTCAGAGGCTCTTTGGAGGCTACCAAGGGGCA	722		Db	1746	TGAGACACAGCTTATGTCTGACTCGCTTCAGGCCAAACAAAGGTGTATCGAGTACGATTTTC	1805	
Db	666	CAGCAGGAGGCGCTACTCAGTGTGGCCCAAGAGGCTCCTGGCGTCCACCAAGGGGCA	725		QY	1803	AGCTGGCACTGGCTGTGCTATGAGTCCCTTCTCAGTGGATGCGACACAGCACTCGG	1862	
QY	723	GGATGTGGTCAATTTGGAGCCCGCAGAGAACACACAGGTAGTGTCTGCGACAGAAATGTAGT	782		Db	1806	GGCTGGTACAGCAGCGGCTTCGGGGCGCCCTCCAGTGGATGCAATGAGGTCAGCGCCAG	1865	
Db	726	GGAGTGGTCAATTTGGAGCCCGCAGAGAACACACAGTGGTGTCTGGCCAGAGTGTGT	785		QY	1863	TGTGCACACACAGAGCCATGTTCCCTTTGGCCCTCGAGNAATTCAGAGGTGAGGGCAAGAT	1922	
QY	783	GATGAGTGGTGGCTGTGCTGACCCCGCCCTTTTGTGTCTGGTTCGAGCAGAGATGG	842		Db	1866	TATGCACAACCAAGAGCCATGTCCTTTTGGCCCTTCGAGAGTTGAGAGGTGCAAGGCAAGAT	1925	
Db	786	GATGGAATGTGTGGCTCAGCTGACCCCGCCCTTTTGTGTCTGGGTCCGACACAGACGG	845		QY	1923	GGAGTCCCTGTGTGTGATGGCAGCGCCCTCTCACCCACCCAGATCTCTGGATACAA	1982	
QY	843	AAAGCTATCTCCACGATGTCTGTTCTGGGCGGACCAATCTACTCATCGCCAGCGC	902		Db	1926	GGAGTCCCTGTGTGTGATGCGACGACCCCTCTCACCCCAACAGATCTCTGGCTACAA	1985	
Db	846	GAAGCCCATCTCCACAGATGTCTGCTGTGGGCGGACCAACTACTAATTCACCAAGCG	905		QY	1983	ACTCTACTGGGAGAGGTGGGAAACAGAGGAGGAGCAGATGTGACCCGCCCCAGGGGG	2042	
QY	903	GCAGCTTCGGCACTCTGGAGTCTATGTCTGCGGAGCCCAACAGCCCTCTCACGCGTGAATT	962		Db	1986	ACTATATTGGCGGAGGTGGGGCTGAGGAGGAGGCCAATGGCGATCGCTCGCAGGGGG	2045	
Db	906	GCAGCCCTGGCACTTCGGGCTCTATGTCTGCGCGGCCAACAGCCCGCGACGCGGACTT	965		QY	2043	TGCTGGAGATCAAGCTTGGGACGTGGGGCCGTGGGCTGAGAGAGAAAGTGAAGCAGTA	2102	
QY	963	CGCCACTGCGGCTGTGAGTCTCGAGTGTGTGTGCCCCAGCCATCTGCGAGGCAACCGA	1022		Db	2046	CGTGGAGACCAAGGCTTGGGATGTGGGGCTCTGCGGCTCAAGAGAAAGTGAAGCAGTA	2105	
Db	966	CGCCACTGCGAGCGCTGAGTCTGCTGTGGGCGCTCCCGCCATCACTCAGCGGCCGA	1025		QY	2103	TGAACTGACCCCAAGTTAGTCCCTGGCAGCGCCGTAACAGGTGAAGCTCTGAGCTTTCAACAA	2162	
QY	1023	GGCGCTTCGCGAGCGGGCCAGCAGCGCGCTTCTGTGTGCGGGCGTTCGCGGAGCC	1082		Db	2106	TGAGCTGACCCAGCTAGTCCCTTGGCGGCTGTACAGAGTGAAGCTCTGGCTTTCAACAA	2165	
Db	1026	GGCGCTGTGCGGACGCGGGCGAGCACAGCGCGCTTCTGTGTGCGCGCTGCGGGAGCC	1085		QY	2163	ACACGAGGACGGCTACGCTGTGTGTGGAAGGGCAAGACGGAGAGGCGCCCAACGCCA	2222	
QY	1083	ACGGCCCGCTGCACTGGCTGCAACAGCGGATCCGTTGCGGACCCAAATGGGGCGCTCAA	1142		Db	2166	ACATGAGGATGGCTATGACAGCAGTGTGGAAGGGCAAGACGGAGAGGCGCCGCGCACGA	2225	
Db	1086	CGGGCCAGCGCTGCTGGCTGCAACAGGGGGCGCTGCTGCGGCCCAACGGGGCGCTCAA	1145		QY	2223	CCTGCTATCCAGAGGGGCCACCGCTGCTCTGCGCCATGTGCCAGAGTCAACAG	2282	
QY	1143	GGTGACGGCGGTGGCGGACGTTGGTCACTCACTAGATCGGCTTCGAGGACGCTGGCTA	1202		Db	2226	CATGCTATCCAGAGGGGACCAACCCCTGCTCCAGCCCACTGTCATGCGGAAATCAACAG	2285	
Db	1146	GGTCCAGGGCGGGTGGCAGCTGGTCAACACAGATCGGCTTCGAGGACCGCGGCTA	1205		QY	2283	CTCCACTTCCATTTGGCTTCGGTGAAGGACAGACTTTTACACCTGTCAAGATTGTCAA	2342	
QY	1203	CTACCACTGAGTACGAGAAACAGCGGGGAACTGCGCTGTGCGCTGCGCCCTGGCGGT	1262		Db	2286	CTCCACATCCATCTGGCTTCGGTGGAAAGGACAGATTTACACACAGTCAAGATTGTCAA	2345	
Db	1206	CTACCACTGAGTGTGCTGAGAAACAGCGGGGAAATGGCGTGGCTGCGCTGCGTGGCGGT	1265		QY	2343	CTACACTGTGACGCTTTCGGCCCTGGGGGCTCAGGAAATGCTTCCCTGGTCACTACTATAC	2402	
QY	1263	AGTGGTGGCGAGGGGCTGCCAGCGCCCGACTCGGGTCAACAGCCAGCGCGCTGAGCAG	1322						

Db 2346 CTACACTGTGCGCTTCAGCGCCCTCGGGGCTCAGGAATGCTCTCCCTGGTCACTATTACAC 2405
 Qy 2403 CAGCTCTGAGAGACATTTCTCATTTGGGGCTGAAACCATTTTACCAAGTACGAGTTTGC 2462
 Db 2406 CAGTTCTGAGAGACATCTCTCATTTGGGGCTTGAAGCAATTCACCAATACGATTTGC 2465
 Qy 2463 GGTACAGTCCACGAGGTGATATGATAGGGCCCTTTTGGCTCGGTGATAGACGCTCCAC 2522
 Db 2466 AGTGCAGTCTACCGGCTGAGACATGATGGGCTTTTGGCTCTGTGTGGAGCGCTCCAC 2525
 Qy 2523 CTGTGCTGACCGGCTTCAACACCTCTCTGACCTGCGCTGAGCCCTGACACCATC 2582
 Db 2526 CTGTGCTGACCGGCTTCAACACCTCTCTGACCTGCGCTGAGCCCTGACACCATC 2585
 Qy 2583 CACCGTTCCGTTACACTGTGTGTCCTCCACGAGGCCCAATGATGATGATGATGATGATCT 2642
 Db 2586 CACGTTCCGCTGCACTGTGTGTCCTCCACGAGGCCCAATGATGATGATGATGATGATCT 2645
 Qy 2643 AATTCCTACAGCAACAGCAACACGAGGCCCAATGATGATGATGATGATGATGATCT 2702
 Db 2646 GATCCTGTACAGCAACAGCAACACGAGGCCCAATGATGATGATGATGATGATGATCT 2705
 Qy 2703 GGGAAACATCTTCACTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGT 2762
 Db 2706 GGGAAACATCTTCACTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGT 2765
 Qy 2763 GATGGAGGCCCAACAGCAACACGAGGCCCAATGATGATGATGATGATGATGATGATCT 2822
 Db 2766 GATGGAGGCCCAACAGCAACACGAGGCCCAATGATGATGATGATGATGATGATGATCT 2825
 Qy 2823 TCTGAAGAGACATTTCTGAGCTCTCTGAGTGTGACGCGCTGACGCGCATCATCGTGG 2882
 Db 2826 GCTCCAGGAGAGCTGTGAGCTCTCTGAGTGTGACGCGCTGACGCGCATCATCGTGG 2885
 Qy 2883 TGTCTGCTGGGCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2942
 Db 2886 TGTCTGCTGGGCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2945
 Qy 2943 CCACAGGAGGCCCTTCCGAGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3002
 Db 2946 CCACAGGAGGCCCTTCCGAGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3005
 Qy 3003 CACAGAGCTCGGCTTGGGCTTCCAGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3062
 Db 3006 CTTCCAGAGCTCGGCTTGGGCTTCCAGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3065
 Qy 3063 GCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3122
 Db 3066 GCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3125
 Qy 3123 ACACAGCTTATGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3182
 Db 3126 GCACAGCTTATGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3185
 Qy 3183 CTGGGCTCAGGCGGAGCAACAACTGGGCGAGGCTCTGGGCGAGGCTGTGAGTCTGCCCA 3242
 Db 3186 CTGGGCTCAGGCGGAGCTGAGTGTGGCTGGTCTCTGGGCGAGGCTGTGAGTCTGCCCA 3245
 Qy 3243 GGGTATGCTCAAGCGGCTGCTGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3302
 Db 3246 GGGTATGCTCAAGCGGCTGCTGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3302
 Qy 3303 GACATGCTGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3362
 Db 3303 GACGCT 3362
 Qy 3363 GTCCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3422
 Db 3363 ACCCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3422
 Qy 3423 ATCCAAAGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3482
 Db 3423 ATCTAACGGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3482

Qy 3483 AGAGACTGTGCTTCCAGCTCTGGAGCTGTGGAATCTGTCTCAAGGAGCAGACTGGCTGGG 3542
 Db 3483 AGAGCTCTCTGATCTCATCTCGGTGTGGGATCCAGGGCAGGGGAGCTGGCTGGA 3542
 Qy 3543 CAGGAGCTGGGAGGGTCCCAACCAACCAACAGTGGGCGCAGAGAGGCTCACCTGCTGGC 3602
 Db 3543 CAGGAGTGTGGGAGGGTGTGAGCTGGCAGCCCGCGGCCAGACAGACTTACCTGCTGGC 3602
 Qy 3603 AGAAGCAGCAGCTGCTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 3662
 Db 3603 AGAGCAGCAGCTGCTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 3662
 Qy 3663 GGGCCCTGGGAAAAGCTGCCAGGCCCAAGCCCTGTGTCTCTTAACAGTCAAGCCCAAGCT 3722
 Db 3663 GACCCCTGGAGATAGCTGCCAGCTCAATCCCTGCTGCTCTAGGAGCCAGCCAGGCT 3722
 Qy 3723 TCCAGAGGCCCTGTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCT 3747
 Db 3723 GCCCAGATCCCGGTCT 3747

RESULT 3

AADI0022

ID AADI0022 standard; cDNA; 2796 BP.

XX AC AADI0022;

XX DT 12-SBP-2001 (first entry)

XX DE Mouse Nope (neighbour of punc ell) extracellular domain cDNA.

KW KW Mouse; Nope; neighbour of punc ell; cytostatic; neuroprotective; vaccine;
 gene therapy; cerebroprotective; colonic cancer; mental retardation;
 tumour suppressor; chromosome 9; transgenic animal; genetic disorder;
 obesity; Bardet-Biedl syndrome; autosomal recessive disorder; retinitis;
 polydactyly; pigmentosa; hypogonadism; hypertension; diabetes mellitus;
 renal anomaly; cardiovascular anomaly; extracellular domain; es.
 OS Mus musculus.

XX FH Key Location/Qualifiers

XX CDS 1..2796

XX PT /*tag= a

XX PT /product= "Mouse Nope extracellular domain"

XX PT /note= "CDS does not include start and stop codon"

XX PT /partial

XX PN WO200149714-A2.

XX PD 12-JUL-2001.

XX PP 26-OCT-2000; 2000WO-US29698.

XX PR 04-JAN-2000; 2000US-0174496.

XX PR 19-MAY-2000; 2000US-0205789.

XX PA (NEUR-) NEUROSCIENCES RES FOUND INC.

XX XX Salbaum JW;

XX XX WPI; 2001-441846/47.

XX XX P-FSDB; AAE05252.

XX DR Murine Nope polypeptides and nucleic acids useful for preventing,

XX DR diagnosing and treating colonic cancer and Bardet-Biedl syndrome -

XX PS Claim 11; Page 77-81; 99pp; English.

XX CC

XX CC The present invention relates to Nope (neighbour of punc ell) which is

XX CC used in the prevention, treatment and diagnosis of diseases associated

XX CC with inappropriate Nope expression such as cancers especially colonic

XX CC cancer and genetic disorders, as Nope is thought to be a tumour

CC suppressor. Nope gene is located on chromosome 9 and is used in gene
CC therapy. Nope is used as vaccine. Nope gene may be administered to treat
CC diseases by rectifying mutations or deletions in a patient's genome that
CC affect the activity of Nope by expressing inactive proteins or to
CC supplement the patients own production of Nope polypeptides. Nope gene
CC is used to study the expression and function of Nope polypeptides and
CC their role in metabolism through the creation of transgenic animal
CC models. The anti-Nope antibodies and Nope antagonists may also be used
CC to down regulate Nope expression and activity for the treatment of
CC Bardet-Biedl syndrome which is an autosomal recessive disorder
CC characterised by mental retardation, obesity, polydactyly, retinitis
CC pigmentosa and hypogonadism. Patients with Bardet-Biedl syndrome have a
CC high incidence of hypertension, diabetes mellitus and renal and
CC cardiovascular anomalies. The present sequence is mouse Nope (neighbour
CC of punc ell) extracellular domain cDNA.

XX Sequence 2796 BP; 591 A; 849 C; 821 G; 535 T; 0 other;

Query Match 45.3%; Score 2796; DB 22; Length 2796;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2796; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 64 GGGAGCTGCCATTGGCCCGGAGGACAACTGTCAAGCTGAGCTGTGTATGAGGGACCCCTG 123
Db 1 GGGAGCTGCCATTGGCCCGGAGGACAACTGTCAAGCTGAGCTGTGTATGAGGGACCCCTG 60
Qy 124 CAAGTATCTTGGCCCTGAGCAGGCTGTGTCTGATGCACTTTGGGGGCTACAGCT 183
Db 61 CAAGTATCTTGGCCCTGAGCAGGCTGTGTCTGATGCACTTTGGGGGCTACAGCT 120
Qy 184 GCTGGGCTCCGACGAGGTTGACATGAGCAAGATGAGACACTGTACTAGAGCATGAG 243
Db 121 GCTGGGCTCCGACGAGGTTGACATGAGCAAGATGAGACACTGTACTAGAGCATGAG 180
Qy 244 AACCTGCACCTGTACCAATGGCTCCCTGTGGCTGTCTCACCCCTAGAGCAAGAGAC 303
Db 181 AACCTGCACCTGTACCAATGGCTCCCTGTGGCTGTCTCACCCCTAGAGCAAGAGAC 240
Qy 304 AGCGATGATGAGGAGCTCTTAGGATCTGAGAGTCACTGAGGGCAGCTATCTGTCTG 363
Db 241 AGCGATGATGAGGAGCTCTTAGGATCTGAGAGTCACTGAGGGCAGCTATCTGTCTG 300
Qy 364 GCCACAGCCGCTAGGAGTGTGGCCGACGAGTGTCTGTGGTCAAGCTTGGCCACACTC 423
Db 301 GCCACAGCCGCTAGGAGTGTGGCCGACGAGTGTCTGTGGTCAAGCTTGGCCACACTC 360
Qy 424 GAAGACTTCTCTGCAACCCGAGTCCAGATTGTGGAGGAGAACGGGACAGCAGCTTT 483
Db 361 GAAGACTTCTCTGCAACCCGAGTCCAGATTGTGGAGGAGAACGGGACAGCAGCTTT 420
Qy 484 GAATGCCACACAAAGGCTTTCCAGCCCGCCATCAATTATTTGGGAAAAGGACAGGTGACC 543
Db 421 GAATGCCACACAAAGGCTTTCCAGCCCGCCATCAATTATTTGGGAAAAGGACAGGTGACC 480
Qy 544 GTGCTGAGGAGCCCGCTCATCACTTCCCAAGTGGCTCTCCAGATCCTAGATGTC 603
Db 481 GTGCTGAGGAGCCCGCTCATCACTTCTCCCAAGTGGCTCTCCAGATCCTAGATGTC 540
Qy 604 CAGGACAGTGTAGGCTCTTACCGCTGTGGCCACCAATTCAGCCCGCCCAAGATTTC 663
Db 541 CAGGACAGTGTAGGCTCTTACCGCTGTGGCCACCAATTCAGCCCGCCCAAGATTTC 600
Qy 664 AGCCAGGAGGCTCGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTACAGGGGGCAG 723
Db 601 AGCCAGGAGGCTCGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTACAGGGGGCAG 660
Qy 724 GATGTGTCTATGTGGAGCCCGCAGAGAACCAACCGTAGTGTCTGGACAAATGTAGTG 783
Db 661 GATGTGTCTATGTGGAGCCCGCAGAGAACCAACCGTAGTGTCTGGACAAATGTAGTG 720
Qy 784 ATGAGTGCCTGGCTCTGTGTAGCCCAACCCCTTTTGTGTCTGGTCCGACAGATGGA 843
Db 721 ATGAGTGCCTGGCTCTGTGTAGCCCAACCCCTTTTGTGTCTGGTCCGACAGATGGA 780

Qy 844 AAGCCTATCTCCACGAGTGTCTGTCGGCCGAGCAACAACTACTCATCTCCAGCGCG 903
Db 781 AAGCCTATCTCCACGAGTGTCTGTCGGCCGAGCAACAACTACTCATCTCCAGCGCG 840
Qy 904 CAGCCTCGGCACTCTGGAGTCTATGTCTGTCGGAGCAACAAAGCCCTCAGCGTGTCTTC 963
Db 841 CAGCCTCGGCACTCTGGAGTCTATGTCTGTCGGAGCAACAAAGCCCTCAGCGTGTCTTC 900
Qy 964 GGCATCGCGCTGTCTGAGCTCCAGTCTGTCGTCGCCCGAGCCATCTCGCAGGACCCGAG 1023
Db 901 GGCATCGCGCTGTCTGAGCTCCAGTCTGTCGTCGCCCGAGCCATCTCGCAGGACCCGAG 960
Qy 1024 GGGCTCTCCGAGCGCGGCGCAGCAGCGCGCTCTGTGTGCGCGGCTCTCCGGGAGCCA 1083
Db 961 GGGCTCTCCGAGCGCGGCGCAGCAGCGCGCTCTGTGTGCGCGGCTCTCCGGGAGCCA 1020
Qy 1084 CCGCCCGCTGTGCACTGGCTGCAAGCGGATTCCTGTTCGAGCCCAATGGGCGCGTCAAG 1143
Db 1021 CCGCCCGCTGTGCACTGGCTGCAAGCGGATTCCTGTTCGAGCCCAATGGGCGCGTCAAG 1080
Qy 1144 GTGCAGGCGGTGTGGCGGAGCTTGTGTATCATCTCAGATCGGCTGTCAGGACGCTGGCTAC 1203
Db 1081 GTGCAGGCGGTGTGGCGGAGCTTGTGTATCATCTCAGATCGGCTGTCAGGACGCTGGCTAC 1140
Qy 1204 TACCAGTGTAGCAGAAAACAGCGCGGAACTGTCTGTGCGCTGCGCCCTCTGGCGGTA 1263
Db 1141 TACCAGTGTAGCAGAAAACAGCGCGGAACTGTCTGTGCGCTGCGCCCTCTGGCGGTA 1200
Qy 1264 GTGCTGCGGAGGGCTGCGCCAGCGCCCGAGCTCGGGTCAAGCCACGCGCTGAGCAGC 1323
Db 1201 GTGCTGCGGAGGGCTGCGCCAGCGCCCGAGCTCGGGTCAAGCCACGCGCTGAGCAGC 1260
Qy 1324 TCCTCTGTGTGTGTGCTGTGGAGCGGCTGTGATGTCACAGCGAGCAAAATCATTTGGCTTC 1383
Db 1261 TCCTCTGTGTGTGTGCTGTGGAGCGGCTGTGATGTCACAGCGAGCAAAATCATTTGGCTTC 1320
Qy 1384 TCTCTTCACTTACCAAAAGGCAAGGAGTGGACAAATGTGTGAGTACCAAGTTTGCAGTAAAC 1443
Db 1321 TCTCTTCACTTACCAAAAGGCAAGGAGTGGACAAATGTGTGAGTACCAAGTTTGCAGTAAAC 1380
Qy 1444 AATGACACCAAGAGCTGCGGAGCTTGGGAGCTTGGAAACCCCAACCGATTTATGATTTCTAC 1503
Db 1381 AATGACACCAAGAGCTGCGGAGCTTGGGAGCTTGGAAACCCCAACCGATTTATGATTTCTAC 1440
Qy 1504 GTGCTGCGCTTACTCCAGCTTGGGGGCGAGCGAACTCTCAGCCAGCCCTGTGTGCTATACA 1563
Db 1441 GTGCTGCGCTTACTCCAGCTTGGGGGCGAGCGAACTCTCAGCCAGCCCTGTGTGCTATACA 1500
Qy 1564 CTGGAGCATGTCTCCAGCGCAGCAACCCAGCTTACCTTGTCTCAGCCCAACCCCTCGGAC 1623
Db 1501 CTGGAGCATGTCTCCAGCGCAGCAACCCAGCTTACCTTGTCTCAGCCCAACCCCTCGGAC 1560
Qy 1624 ATCAGGCTGTCATGGCTGCGCTTCCAGCTTCCAGCTTCCAGCAATGGAGTGTGAAATTC 1683
Db 1561 ATCAGGCTGTCATGGCTGCGCTTCCAGCTTCCAGCTTCCAGCAATGGAGTGTGAAATTC 1620
Qy 1684 AAGATAGATGATGAGTTTGGGAAAGGAGATCAGGTTTCTCCACCGAGGTCCTGGAAAT 1743
Db 1621 AAGATAGATGATGAGTTTGGGAAAGGAGATCAGGTTTCTCCACCGAGGTCCTGGAAAT 1680
Qy 1744 GAGACAACTTACGTTTAACTCACTTTCAGCCAAACAAAGTGTACCGAGTCCGGAATTTCA 1803
Db 1681 GAGACAACTTACGTTTAACTCACTTTCAGCCAAACAAAGTGTACCGAGTCCGGAATTTCA 1740
Qy 1804 GCTGGCACTGGCGCTGGCTATGAGTCCCTTCTCAGTGTGATGACAGCAGGACACCTGGT 1863
Db 1741 GCTGGCACTGGCGCTGGCTATGAGTCCCTTCTCAGTGTGATGACAGCAGGACACCTGGT 1800
Qy 1864 GTGCACAAACAGAGCATGTTCCCTTTGCGCTTGTAGAGTTGAGGTCGAGGCAAGATG 1923
Db 1801 GTGCACAAACAGAGCATGTTCCCTTTGCGCTTGTAGAGTTGAGGTCGAGGCAAGATG 1860

QY 63 CGGGAGCTGCCATTGCCCCCAGGAGACAACTGTCAAGCTGAGCTGTGATGAGGACCCCT 122
DB 69 CGGGAGCTGCTGTTGCCCCCAGGAGACGACTGTGGAGCTGAGCTGTGGAGTGGGGCACT 128
QY 123 GCAAGTGAATCCTGGGCCCCCTGAGCAGGCTGTGGTGTGGAATGCACTTTGGGGGCTACAGC 182
DB 129 GCAAGTGAATCCTGGGCCCCCAGGAGGCTGCACTGTAACTGTAGCCCTGGGGGCTGCTGC 188
QY 183 TGCTGGGCTCCGACAGGCTGACATGGAGCAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 242
DB 189 CGCTGGAGCCCCCAGGAGGCTGACCTGGAGCAAGGATGGGGACACCCCTGTGGAGCAGCA 248
QY 243 GAACCTGACCTGTAACCAATGCTCCCTGTGGCTGTCTCAACCCCTAGAGCAAGAGGA 302
DB 249 CCACTTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 308
QY 303 CAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGTCACTGAGGGGAGCTATTCCTGTCT 362
DB 309 CAG---TGAGGAGTCACTCCCTGAGGCTGTGGGGTCAATTAAGGCAACTATTCGTGCT 365
QY 363 GGGCCACAGCCCGCTAGGAGTGGTGGCCAGCCAGGTTGCTGTGGTCAAGCTTGGCACT 422
DB 366 AGCCACAGCCCGCTCGAGTGTGGCCAGCCAGACTGCTGTGTCAAGCTTGGCACT 425
QY 423 CGAAGACTTCTCTGTCACCCCGAGTCCAGATTTGTGGAGGAGAAAGGAGCAGCAGCTT 482
DB 426 CGCAGACTTCTCTGTCACCCCGAGTCTCAGACGGTGGAGGAGAAAGGAGCAGCTGCTT 485
QY 483 TGAATGCCACACCAAGGGCCCTTCAGCCGCCATCAATTACTTGGGAAAGGACCAAGTGC 542
DB 486 TGAATGCCACACCAAGGGCCCTTCAGCCGCCATCAATTACTTGGGAAAGGACCAAGTGC 545
QY 543 CGTGCCTGAGGAGCCCGCTCATCACTCTTCCCAAGTGGCTCTCCAGATCTCTAGATGT 602
DB 546 ATTGCCTGAGGAGCCCTCGGCTCATCGTGTCTTCCCAAGGCGGTCTTCCAGATCTCTAGATGT 605
QY 603 CCAGGAGTGAATGAGGCTCTTACCGCTGCTGGCCACCAATTAAGCCCGCCCAACGAT 662
DB 606 TCAGGAGTGAATGAGGCTCTTACCGCTGCTGGCCACCAATTAAGCCCGCCCAACGAT 665
QY 663 CAGCCAGGAGCCCTGCTCATCTGTGGCCCTCAGAGGCTCTTGGAGGCTACAGAGGGGCA 722
DB 666 CAGCCAGGAGCCCTGCTCATCTGTGGCCCTCAGAGGCTCTTGGAGGCTACAGAGGGGCA 725
QY 723 GGAATGGTCAATTTGGAGCCGCCAGAGAACACCAAGTGTGTCTGGACAGATGTAGT 782
DB 726 GGACGTGGTCAATTTGGAGCCGCCAGAGAACACCAAGTGTGTCTGGACAGATGTAGT 785
QY 783 GATGGAGTGTGGGCTCAGCTGACCCCAAGCCCTTTTGTGTCTGGGTCGAAGATGG 842
DB 786 GATGGAGTGTGGGCTCAGCTGACCCCAAGCCCTTTTGTGTCTGGGTCGAAGATGG 842
QY 843 AAAGCCTATCTCCACGATGTCACTGTGGCCCGGACCAATCACTATCGCCAGCGC 902
DB 843 GAAGCCATCTCCACGATGTCACTGTGGCCCGGACCAATCACTATCGCCAGCGC 902
QY 903 GCAGCTCGGACCTCTGAGTCTATGTCTGGCCAGCCCAAGCCCTCTCAGCGTACTT 962
DB 903 GCAGCTCGGACCTCTGAGTCTATGTCTGGCCAGCCCAAGCCCTCTCAGCGTACTT 962
QY 963 CGCCACTCGGCTCTGAGCTCCAGTGTCTT---GCTGCCCGCCAGCCATCTCGAGGCA 1019
DB 963 CGCCACTCGGCTCTGAGCTCCAGTGTCTT---GCTGCCCGCCAGCCATCTCGAGGCA 1022
QY 1020 CGAGCGCTCTCGGAGACCGGGGCGAGCAGCGCGCTTGTGTGGCCGGGCTCGGGGA 1079
DB 1023 CGAGCGCTCTCGGAGACCGGGGCGAGCAGCGCGCTTGTGTGGCCGGGCTCGGGGA 1082
QY 1080 GCCACGCGCGCTGCTGAGCTGAGCTGAGCGGATCGCGTGTGGACCGCAATGGGCGCT 1139
DB 1083 GCGCGCGCGCTGAGCTGAGCTGAGCGGATCGCGTGTGGACCGCAATGGGCGCT 1142
QY 1140 CAAGGTGAGGCGGCTGGCGGAGCTTGGTCACTCACTGAGTGGCTCGAGGAGCGCTGG 1199

DB 1143 CAAGGTGAGGCGGCTGGCGGAGCTTGGTCACTCAACAGATCGGCTCGAGGACCGCG 1202
QY 1200 CTACTACAGTGGTAGCAGAAACAGCGGGGAGTGTGCTGTGGCTGTGGCCCTGTGC 1259
DB 1203 CTACTACAGTGGTAGCAGAAACAGCGGGGAGTGTGCTGTGGCTGTGGCCCTGTGC 1262
QY 1260 GGTAGTGGTGTGGCGAGGGGCTGCCCCAGCGCCCGACTCGGGTCAACAGCAGCGCGCTGAG 1319
DB 1263 GGTAGTGGTGTGGCGAGGGGCTGCCCCAGCGCCCGACTCGGGTCAACAGCAGCGCGCTGAG 1322
QY 1320 CAGCTCTCTGTGTGTGGCTGTGGAGCGGCTGTAGTTGCAAGCGAGCAAAATCAATGG 1379
DB 1323 CAGCTCGGCTGTGTGTGGCTGTGGAGCGGCTGTGGAGCGGCTGTGGCTGTGGCTGTGG 1382
QY 1380 CTTCTCTCTCTACTACCAAAAGGAGGAGTGGAGCAATGTGGATCAAGTGTGGCT 1439
DB 1383 CTTCTCTCTCTACTACCAAAAGGAGGAGTGGAGCAATGTGGATCAAGTGTGGCT 1442
QY 1440 AAACATGACACACAGAGCTGAGGTTGGGAGCTGTGGAACTGTGGAACTGTGGATTTAGTGT 1499
DB 1443 GAACAAAGCAGCAGCAGAACTACAGGTTGGGAGCTGTGGAACTGTGGATTTAGTGT 1502
QY 1500 CTAAGTGTGGCTACTCTCCAGCTGGGGGCTGAGCGGAGCTCTCAGCGCCAGCGCTGTGCA 1559
DB 1503 CTAGCTGTGGCTACTCTCCAGCTGGGGGCTGAGCGGAGCTCTCAGCGCCAGCGCTGTGCA 1562
QY 1560 TACACTGGAGGATGTCTCCAGCGCAGCAGCGGCTGTAGCTGTGCTGAGCGCCAGCGCTC 1619
DB 1563 CACACTGGATGTGTCTCCAGCTGAGCAGCGGCTGTGCTGAGCGCCAGCGCTGTGCA 1622
QY 1620 GGAATCAGGCTGGAGTGTCTCCAGCTGTGGGAGGAGATCAGGTTTCTCAGCGAGTGTG 1679
DB 1623 GGAATCAGGCTGGAGTGTCTCCAGCTGTGGGAGGAGATCAGGTTTCTCAGCGAGTGTG 1682
QY 1680 GTAAGATGAGTACGCTTGGGAGGAGATCAGGTTTCTCAGCGAGTGTGCTG 1739
DB 1683 GTAAAGATGAGTACGCTTGGGAGGAGATCAGGTTTCTCAGCGAGTGTGCTG 1742
QY 1740 AAATGAGACAACTTACGTTAACTCACTTACGCGGAGGAGTGTGCTGAGTGTGCTG 1799
DB 1743 AAATGAGACAACTTACGTTAACTCACTTACGCGGAGGAGTGTGCTGAGTGTGCTG 1802
QY 1800 TTGAGTGTGGCTGTGGCTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGT 1859
DB 1803 TTGAGTGTGGCTGTGGCTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGT 1862
QY 1860 TGGTGTGCAACACAGAGCAGTGTCTTGTGGCTGTGGAGTGTGGAGTGTGGAGTGTGGAGT 1919
DB 1863 CAGTATGCAACACAGAGCAGTGTCTTGTGGCTGTGGAGTGTGGAGTGTGGAGTGTGGAGT 1922
QY 1920 GATGAGTGTGGCTGTGGCTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGT 1979
DB 1923 GATGAGTGTGGCTGTGGCTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGT 1982
QY 1980 CAACTCTACTGGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 2039
DB 1983 CAACTCTACTGGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 2042
QY 2040 GGGTGTGGAGTCAAGCTTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGT 2099
DB 2043 GGGTGTGGAGTCAAGCTTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGT 2102
QY 2100 GTATGAACTGACCCAGTGTGGCTGTGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGT 2159
DB 2103 GTATGAACTGACCCAGTGTGGCTGTGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGT 2162
QY 2160 CAAACAGGAGTGTGGCTGTGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGT 2219
DB 2163 CAAACAGGAGTGTGGCTGTGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGT 2222
QY 2220 AGACCTGCTGTATCAGAGGGGGGCGAGCGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 2279

Db 2223 AGACATGCTATCCAGAGGGGACACCCCTGCTCCAGCCACAGTCCATGCGGAATCAA 2282
QY 2280 CAGCTCCACTTCCATTGCTTGGTGGAGAGCCAGACTTTACCACTGTCAAGATTGT 2339
Db 2283 CAGCTCCACTTCCATTGCTTGGTGGAGAGCCAGACTTTACCACTGTCAAGATTGT 2342
QY 2340 CAACTACACTGTACGCTTGGGCTTGGGCTCAGGAATGCTTCCCTGGTCACTACTA 2399
Db 2343 CAACTACACTGTGCGCTTCAAGCCCTTGGGCTCAGGAATGCTTCCCTGGTCACTACTA 2402
QY 2400 TACCAGCTCTGAGAGAGACTTCTCATTTGGCGGCTGAAACCAATTTACCAAGTACAGATT 2459
Db 2403 ---CAGTTCTGAGAGAGACTTCTCATTTGGCGGCTTGAAGCCATTTACCAATATCAGATT 2459
QY 2460 TCGGTACAGTCCCAAGAGTGGATATGATGAGGCTTGTGGCTCGTGTGAGAACGCTC 2519
Db 2460 TCGAGTGCAGTCTCAGCGGCTGACATGATGGGCTTTTCGGCTCTGTGGTGGAGCGCTC 2519
QY 2520 CACCTGCTGACCGGCTTCAACACCTCTTCAACCTGCTGACCTGGCTGAGCCCTGACACC 2579
Db 2520 CACCTGCTGACCGCTTCAACACCTCTTCAACCTGCTGACCTGGCTGAGCCCTGACACC 2579
QY 2580 ATCCACCGTTCCGTTTACACTGCTGTCCTCCACGAGGCCCAATGCTGAGATTGTGAGTA 2639
Db 2580 GTCCACCGTTCCGTTTACACTGCTGTCCTCCACGAGGCCCAATGCTGAGATTGTGAGTA 2639
QY 2640 TCTAATTTCTACAGCAACCAACACACGAGCCCGAACAACAGTGGACCTGCTCACCAC 2699
Db 2640 TCTGATCTCTACAGCAACCAACACGAGCCCTGAGCACCACTGCTCACCAC 2699
QY 2700 -----AGAGGAAACATCTTCAGTGCAGAGTCCATGCTGCTAGAGTGCACCTCGTA 2753
Db 2700 GCAGGCTGAGGAAACATCTTCAGTGCAGAGTCCATGCTGCTGAGAGCGACACTCGGTA 2759
QY 2754 TTTCTTCAAGATGGAGCCCGCACAGAGTGGGCTTGGGCTTTTCGGCTTCCAGTA 2813
Db 2760 CTTCCTCAAGATGGGCGCGCACAGAGTGGGACTTGGGCTTTCTCCGCTCGAGTA 2819
QY 2814 TGTGATTAATCTGCAAGAGACATCTTCAGACTTCTTGGATGTGCACGCGCTCACGGGAT 2873
Db 2820 TGTGATCACTGCTCAGGAGAACTGTTCAGACTGCTGAGCATGCTCAGTCAAGGAT 2879
QY 2874 CATGTGGTGTGCTGCTGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2933
Db 2880 CATGTGGTGTGCTGCTGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2939
QY 2934 ACAAGCTCCACAGGAGCCCTTCCGATGTGCTCCTCAGGACCCAGGAAACCC 2993
Db 2940 CCGAGCCCTCCACAGGAAATCTCTCCAGGCTGCTCCTCCAGCCACCCCGGAAATCC 2999
QY 2994 AGCGCTCTACCAAGAGCTCGGCTTGGGCTTCCAGTGTCTGCTGCTGCTGCTGCTGCTGCT 3053
Db 3000 CGCGCTGTACTCCAGAGTCTGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 3059
QY 3054 GTCCCTCGTGCATCTCTGCTCCAGGATGCTTCCCAACCACTCAGATGTGGAAGACAA 3113
Db 3060 GTCCCTTGTGACCCCTTCCAGGACTGTGCTCCGCGCTCAGAGCTGAGGACAG 3119
QY 3114 GGCTGAGTACACAGCTTATGGTGGAGTGTGCTTCCAGATGCTGCTGCTGCTGCTGCTGCT 3173
Db 3120 GGCTGAGTACACAGCTTATGGTGGAGTGTGCTTCCAGATGCTGCTGCTGCTGCTGCTGCT 3179
QY 3174 AAAAGATCTCTGGGCTCAGGAGGGGACCAACTGGGAGGCTCTCTGGGAGGCTGTGA 3233
Db 3180 AAAGGT---AAGTGTCTAACCAAGCGGCTGAGCTGGGCTGGTCTCTGGGAGGCTGTGA 3236
QY 3234 GCTGCCCAAGGATGTGCTCAAGGCGGCTCTGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 3293
Db 3237 GCTGCCCAAGG---CAGGCTTCCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3293
QY 3294 AACCGGCGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3353
Db 3294 AACTGGCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3347

QY 3354 AAAGAGCGCTTCCAGAGCTGCGAGGAATCAGGTGGAAGTCAATGTTCCTCCGA 3413
Db 3348 GAAGAAAGTCAACCCCGAGCTGCGAGGAACCAAGTGGAGGCTCAAGTCAATGTTCCTCCGA 3407
QY 3414 CTTGCTGCTCAACAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3473
Db 3408 CTTTGTGCTCAACAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3467
QY 3474 ACTGACTGCGAGAGCTTCTGCTTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3533
Db 3468 CTTGCTTCCAGAGGCTCTGATCTCATCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3527
QY 3534 CTGCT 3593
Db 3528 CTGCT 3587
QY 3594 CTGCTTCCAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3653
Db 3588 CTGCTTCCAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3647
QY 3654 TATAGAGAGGCTTCCAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3713
Db 3648 GCTAGAGAGAGCTTCCAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3707
QY 3714 CCCAGGCTTCCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3747
Db 3708 CCCAGGCTTCCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3741

RESULT 5

ABZ11219

ID ABZ11219 standard; cDNA; 3450 BP.

XX AC ABZ11219;

XX XX 20-JAN-2003 (first entry)

XX Human polynucleotide SEQ ID NO 101.

Human; genome mapping; gene therapy; food supplement; virus; fungus;
cell-proliferative disorder; neurodegenerative disease; bacterial;
Parkinson's disease; Alzheimer's disease; autoimmune disease;
multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
arthritis; cytostatic; immunomodulator; neoplastic; dermatological;
anti-parkinsonian; antidiabetic; immunosuppressive; dermatological;
haemostatic; vulnary; fungicide; antibacterial; virucide; protozoicide;
antiarthritic; gene; ss.

XX Homo sapiens.

OS WO200270539-A2.

XX 12-SEP-2002.

XX 05-MAR-2002; 2002WO-US05095.

XX 05-MAR-2001; 2001US-0799451.

XX (HYSB-) HYSBQ INC.

PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;

DR WPI; 2002-759812/82.

XX P-PSDB; ABP69002.

PT New polynucleotides comprising sequences assembled from expressed
sequence tags (ESTs), useful for treating cell-proliferative, or
neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or
platelet or coagulation disorders

XX

PS Claim 1; SEQ ID NO 101; 1012pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (I) comprising a

CC nucleotide sequence selected from any of 948 sequences

CC (ABZ1119-ABZ12066) or their mature protein coding portion, active domain

CC coding protein or complementary sequences. The polynucleotides are useful

CC for identifying expressed genes or for physical mapping of human genome.

CC The encoded polypeptides (ABP68902-ABP69849) are useful as molecular

CC weight markers, as a food supplement, for generating antibodies, in

CC medical imaging, screening and diagnostic assays and for treating

CC cell-proliferative disorders (cancer), neurodegenerative diseases

CC (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple

CC sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid

CC disorders, platelet or coagulation disorders, wound, burns, incision,

CC ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,

CC parasitic), arthritis, etc.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 3450 BP; 681 A; 1050 C; 1067 G; 652 T; 0 other;

Query Match 34.3%; Score 2118.8; DB 24; Length 3450;

Best Local Similarity 80.8%; Pred. No. 0;

Matches 2599; Conservative 0; Mismatches 567; Indels 50; Gaps 9;

991 CTTGCTGCCAGCCATCTCGAGGACACCGAGCGCTCTCGCGAGCGCGGCGAGCAC 1050

DB 184 CTAGCGGCTCCGCCATCACTCAGGCGCGCGAGCGCTCTCGCGAGCGCGGCGAGCAC 243

QY 1051 GCGCGCTCTGTCGCGCGCGCTCCGGGAGCCAGCGCGCGCTGCACTGGCTGCAAC 1110

DB 244 GCGCGCTCTGTCGCGCGCGCTCCGGGAGCGCGCGCGCTGCACTGGCTGCAAC 303

QY 1111 GGGATCCGTTGGACCCCAATGGCGCGCTCAAGGTCCAGGCGCGCGCTGGTGC 1170

DB 304 GGGCGCGCGCTGCGCGCGCTCAAGGTCCAGGCGCGCGCTGGTGC 363

QY 1171 ATCACTCAGATCGGCTGAGGACGCTGGTCTACTACAGTGCCTAGAGAAACAGCGCG 1230

DB 364 ATCACTCAGATCGGCTGAGGACGCGCGCTACTACAGTGCCTGGTGCAGAAACAGCGCG 423

QY 1231 GGAATGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCG 1290

DB 424 GGAATGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCG 483

QY 1291 CCAGCTCGGCTCAGAGCAGCGCGCTGAGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1350

DB 484 CCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543

QY 1351 CTTGAGTTGACAGGAGCAATCAATGCTTCTCTTCTCACTACCAAAAGGCAAGGGA 1410

DB 544 CCGGAGATGACAGGAGCAGATCATCGGCTTCTCTCTCACTACCAAGGCAAGGGA 603

QY 1411 GTGCAATGCTGAGTACCAATGCTGAGTAAACATGACACACAGAGTGCAGGTTGG 1470

DB 604 ATGGAATGCTGAGTACCAATGCTGAGTAAACATGACACACAGAGTGCAGGTTGG 663

QY 1471 GACTTGAACCCCAACAGGATTAAGTCTTACGTTGCTGCTTCTCTCTCACTACCAAGGCAAGGGA 1530

DB 664 GACTTGAACCCCAACAGGATTAAGTCTTACGTTGCTGCTTCTCTCTCACTACCAAGGCAAGGGA 723

QY 1531 AGCGCAACTCCAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1590

DB 724 AGCGCAACTCCAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 783

QY 1591 CAGCTTACCTTGTCCAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1650

DB 784 CAGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 843

QY 1651 TCCAGCTGAGCAATGAGAGGCTGCTGAGTACAGATGAGTACAGTACAGTACAGTACAGTACAG 1710

DB 844 CCAAGCTGAGCAATGAGAGGCTGCTGAGTACAGATGAGTACAGTACAGTACAGTACAGTACAG 903

QY 1711 GATCAGGTTTTCTCCACCGAGTGTCTGGAATGAGACACAACTTACGTTAACTCAGCTT 1770

DB 904 GATCAGATTTTTCTCTACTGAGTGTGAGGAATGAGACACAGCTTATGCTGAATCGCTT 963

QY 1771 CAGCCAAACAAAGTGTACCGAGTCCGGATTTTCAGCTGGCAGCTGGCGCTGATGAGATC 1830

DB 964 CAGCCAAACAAAGTGTATCGAGTACGGATTTTCGGCTGGTACAGCAGCGGCTTCGGGGCC 1023

QY 1831 CTTTCTCAGTGTAGTACAGCAGAGGACACTGCTGTGTGCAACACAGAGCATGTTCCCTTT 1890

DB 1024 CTTTCTCAGTGTAGTACATCAGAGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1083

QY 1891 GCGCTGCGAGATTCAGGCTGAGGCAAGATGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1950

DB 1084 GCGCTGCGAGATTCAGGCTGAGGCAAGATGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1143

QY 1951 CCGCTTCAACCCCAACAGATCTCTGAGTACAACTCTACTTGGGAGAGTGGGAAACAGAG 2010

DB 1144 CCGCTTCAACCCCAACAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1203

QY 2011 GAGGAGGAGATGAGTACCG 2070

DB 1204 GAGGAGGCGCAATGCGGATCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1263

QY 2071 CCGGTCGCGCTCAAGAGAAAGTGAAGCAGTATGAATGAGTCCCTGCTGCTGCTGCTGCTGCTGCTG 2130

DB 1264 CCGGTCGCGCTCAAGAGAAAGTGAAGCAGTATGAATGAGTCCCTGCTGCTGCTGCTGCTGCTGCTG 1323

QY 2131 CCGTACGAGTGAAGCTCGTACTCTTCAACAAACAGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTG 2190

DB 1324 CCGTACGAGTGAAGCTCGTACTCTTCAACAAACAGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTG 1383

QY 2191 AAGGCAAGAGCGGAGAGCG 2250

DB 1384 AAGGCAAGAGCGGAGAGCG 1443

QY 2251 CCGTCTGCGCGCTCAAGAGTGAAGCAGTATGAATGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 2310

DB 1444 CCGTCTGCGCGCTCAAGAGTGAAGCAGTATGAATGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 1503

QY 2311 AAGCCAGATTTTACCACTGCTCAAGATTTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2370

DB 1504 AAGCCAGATTTTACCACTGCTCAAGATTTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1563

QY 2371 CTTAGGAATGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2430

DB 1564 CTTAGGAATGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1623

QY 2431 GCGCTGAAACCAATTTACCAAGTACGAGTTCGGGTACGTTCCCGAGGAGTATGGAT 2490

DB 1624 GCGCTGAAACCAATTTACCAAGTACGAGTTCGGGTACGTTCCCGAGGAGTATGGAT 1683

QY 2491 GCGGCTTGGGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2550

DB 1684 GCGGCTTGGGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1743

QY 2551 TCTGACCTGCGGCTGAGCG 2610

DB 1744 TCGGACCTGCGGCTGAGCG 1803

QY 2611 ACGGAGCCCAATGCTGAGATTTGAGGATCTTAATTTCTTACAGCAACCAACCAACCAACCA 2670

DB 1804 ACGGAGCCCAATGCTGAGATTTGAGGATCTTAATTTCTTACAGCAACCAACCAACCAACCA 1863

QY 2671 CCGGAACACCAAGTGGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2730

DB 1864 CCGGAACACCAAGTGGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1923

QY 2731 GCGCTAGAGATGACACTCGGTTATTTCTTCAAGATGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 2790

DB 1924 GCGCTAGAGATGACACTCGGTTATTTCTTCAAGATGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1983

methodology. The polypeptides are useful in the treatment of disease such as diabetes, breast-, prostate-, colon cancer and other malignant tumors, hyper- and hypotension, obesity, bulimia, anorexia, growth abnormalities, asthma, manic depression, dementia, delirium, mental retardation, Huntington's disease, Tourette's syndrome, schizophrenia, growth, mental or sexual development disorders, and dysfunctions of the blood cascade system including those leading to stroke. The polynucleotides may be used as diagnostic reagents through detecting mutations in the associated gene and for chromosome localization and for tissue expression studies. The polypeptides and polynucleotides may also be used as vaccines.

SQ Sequence 2223 BP; 440 A; 660 C; 735 G; 388 T; 0 other;

Query Match	24.7%;	Score 1522.4;	DB 22;	Length 2223;
Best Local Similarity	82.8%;	Pred. No. 0;		
Matches 1786; Conservative		0; Mismatches 321;	Indels 51;	Gaps 2;

Qy	64	GGGGAGCTGCGCATTTGCCCCAGAGAGACAATGTCTCAAGCTGAGCTGTGTGATGAGGAGACCCCTG	123
Db	76	GGGGAGCTGCTGTGTGTGCCCCAGAGAGACGACTGTGGAGCTGAGCTGTGTGAGTGTGGGGCCACTG	135
Qy	124	CAAGTGATCTCTGGGCCCTGAGCAGAGCTGTGTGTCTGGACTGCACTTTTGGGGGCTTACAGCT	183
Db	136	CAAGTGATCTGTGGGCCAGAGCAGGCTTGAAGTCTAACTGTAGCTTGGGGGCTGTCTGCC	195
Qy	184	GCTGGGCTCCGACCAAGGCTGACATGGAGCAAGGATGGAGACATGTACTAGAGCATGAG	243
Db	196	GCTGGAGCCCCCACCAGGCTGACTGGAGCAAGGATGGGGACACCTGCTGGAGCAAGAC	255
Qy	244	AACCTGCACCTGTATACCCAAATGGCTCCTGTGTGGCTCTCTCAACCTTAGAGCAAGAAGAC	303
Db	256	CACCTTACACCTGTGCCCAATGGTTCCCTGTGTGGCTGCCAGCCACTAGCACCAATGGC	315
Qy	304	AGCGATGATAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCCTGTCTG	363
Db	316	AG---TGAAGAGTCAGTCCCTGAGGCTGTGGGGTCAATTGAGGCACTAATTCGTGCCTA	372
Qy	364	GCCCAAGCCCGTATAGGAGTGTGCCAGCAGAGTGTGTGTGTGTCAGAGTTTGCGCACTC	423
Db	373	GCCCAAGCCCGCTGTGAGTGTGCCAGCAGAGTGTGTGTGTGTCAGAGTTTGGCACACTC	432
Qy	424	GAAGACTTCTCTGCACCCCGAGTCCAGAGTTGTGGAGGAGAACGGGACAGCACCGCTTT	483
Db	433	GCAGACTTCTCTGCACCCCGAGTCTCAGACGGTGGAGGAGAACGGGACAGCTCGCTTT	492
Qy	484	GATGCCACACCAAGGGCTTCCAGCCCCCATCATTTACTTGGGAAAAAGGACAGAGTGACC	543
Db	493	GAGTGCCACATTGAAGGGCTGCCAGCTCCCATCATTTACTTGGGAGAGGACCAAGTGACA	552
Qy	544	GTGCTTGAGGAGCCCGGCTCATCACTCTTCCCAAGTGGCTCTCCAGATCTTAGATGTC	603
Db	553	TTGCTCTGAGGAGCCTGGGCTCATGTGTCTCCCAACGGCGTCTTTCAGATCTTGGATGTT	612
Qy	604	CAGGACAGTGATGCAGGCTCTTACCGCTGCGTGGCCACCAATTCAGCCGCCCAACGATTC	663
Db	613	CAGGAGAGTGATGCAGGCCCTTACCGCTGCGTGGCCCAACCACTCAGTCTGCCAGCACTTC	672
Qy	664	AGCCAGAGGCGCTCGTCTACTGTGGCCCTCAGAGGGTCTTTTGGAGGCTACAGGGGGCAG	723
Db	673	AGCCAGAGGCGCTACTCAGTGTGGCCCAAGAGGGTCCCTGGCGTCCACAGGGGGCAG	732
Qy	724	GATGTGTCATTTGTGGCAGCCCAAGAGAACCAAGGTAGTGTCTGCAAGAGATGTAGTC	783
Db	733	GAGTGGTCAATTTGTGGCAGCCCAAGAGAACCAACAGTGTGTCTTGGCCAGAGTGTGGTG	792
Qy	784	ATGAGTGTGCTGGCTCTGTGACCCCAACCCCTTTTGTGTCTGGGTCCGACAGATGGA	843
Db	793	ATGGAATGTGTGCTCTAGCTGACCCCAACCCCTTTTGTGTCTGGGTCCGACAGACGG	852
Qy	844	AAGCCTATCTCAAGGATGTCAATGTTCTGGGGCGGACCAATCTACTATCTGCCAGCGG	903
Db	853	AAGCCCATCTCCACAGATGTCAATGTTCTGGGGCGGACCAACCTACTAAATTCGCAACGG	912

RESULT 8
AAC79893
ID AAC79893 standard; cDNA; 680 BP.
XX
AC AAC79893;
XX
DT 09-FEB-2001 (first entry)
XX
XX Human secreted protein encoding cDNA for gene 45.
XX
XX Human; secreted protein; cytosolic; antiarthritic; antiasthmatic;
KW immunosuppressive; antiarteriosclerotic; antiinflammatory; nootropic;
KW neuroprotective; antiarrhythmic; tranquiliser; vulnerary; antibacterial;
KW antipsoriatic; antiasthmatic; antiarthritis; antiarthritis; anti-HIV;
KW autoimmune disorder; allergic condition; cardiovascular disorder;
KW cancer; neurological disease; tissue repair; ss.
XX
OS Homo sapiens.
XX
XX W0200055176-A2.
XX
XX 21-SEP-2000.
XX
XX 09-MAR-2000; 2000WO-US06057.
XX
XX 12-MAR-1999; 99US-0124142.
XX
XX 11-JUN-1999; 99US-0138597.
XX
XX 03-DEC-1999; 99US-0168666.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Komatsoulis G;
XX
XX WPI; 2000-638176/61.
XX
XX P-PSDB; AAB44874.
XX
XX Novel 49 human secreted proteins useful for diagnosis, prevention and
PT treatment of disorders including neurological, cell proliferative,
PT cardiovascular, and autoimmune/inflammatory disorders and microbial
PT infections -
XX
XX Claim 1a; Page 358; 405pp; English.
XX
XX This invention describes a novel isolated polypeptide (I) comprising an
CC amino acid sequence at least 95 % identical to a polypeptide sequence
CC selected from 49 polypeptides encoded by polynucleotide sequences
CC included in American Type Culture Collection (ATCC) deposit number
CC 203917, defined in the specification. The products of the invention have
CC cytosolic, antiarthritic, antiasthmatic, immunosuppressive, nootropic,
CC antiarteriosclerotic, antiinflammatory, neuroprotective, antidiabetic,
CC tranquiliser, vulnerary, antibacterial, antipsoriatic, antiarrhythmic,
CC antineumatic, cardiant and anti-HIV activity. (I) or a nucleic acid (II)
CC encoding (I) is useful for preventing, treating or ameliorating a medical
CC condition and for diagnosing a pathological condition or susceptibility
CC to the condition. (I) is useful for identifying a binding partner which
CC affects the activity of the polypeptide and for identifying an activity
CC in a biological sample. (I), (II) or an antibody (IV) specific to (I) is
CC also useful for treating or preventing a disease, disorder or condition
CC associated with aberrant expression of (I). Diseases treated or diagnosed
CC include immune disorders such as autoimmune diseases, blood protein
CC disorders, anemia, allergic reactions and conditions such as asthma,
CC organ rejection or graft-versus-host disease, inflammation, hyper
CC proliferative disorders, cardiovascular disorders such as arterioarterial
CC fistula, arrhythmias, arteriosclerosis, coronary thrombosis, organ
CC regeneration, cancer, neovascular glaucoma, diabetic retinopathy,
CC rheumatoid arthritis, psoriasis, diseases associated with increased
CC apoptosis that include acquired immunodeficiency syndrome (AIDS),
CC neurological diseases such as Parkinson's disease, viral, bacterial,
CC fungal or parasitic diseases. They are also used to repair, replace or
CC protect tissue damage by congenital defects, to treat trauma, in surgery,
CC including cosmetic plastic surgery, to treat fibrosis, reperfusion injury
CC or systemic cytokine damage, to stimulate chondrocyte growth, to prevent
CC skin aging due to sunburn, to change a mammal's mental state or physical

CC state by influencing biorhythms, cardiac rhythms, depression, memory,
CC stress and for accelerating wound healing. (I), (II) and/or their agonist
CC or antagonist are useful as food additives or preservatives to increase
CC or decrease storage capabilities, fat content, lipid, protein,
CC carbohydrate, vitamin, mineral or other nutritional components. (I) is
CC useful for screening therapeutic compounds. (II) is useful in forensic
CC biology for detecting DNA sequences and as diagnostic probes for
CC detecting the presence of specific mRNA in a particular cell type.
XX
SQ Sequence 680 BP; 251 A; 97 C; 100 G; 232 T; 0 other;

Query Match 5.5%; Score 337.2; DB 21; Length 680;
Best Local Similarity 79.9%; Pred. No. 5.5e-77;
Matches 473; Conservative 0; Mismatches 108; Indels 11; Gaps 6;

QY 5588 GAATGAAATAACCGTTTAAACAGTAGGCTCTTAGCATCACACATAGCTCTCTCAT 5647
DB 11 GAATTAATGGCTGTGTTTAAACAGTAGGCTCTTAGCATTAATACCATAGCTCTCTCAT 70

QY 5648 GTTCTTGTGTTTAAACAGCACTTTCAGGTTCTGGGTTTAAATTAATAGTGCATGAGACA 5707
DB 71 GTTCTTGTGTTTAAACAGCACTTTCAGGTTCTGGGTTTAAATTAATAGTGCATGAGACA 128

QY 5708 ATTTATACCCATTTAGGCTGGGTGGAATTTG-TTCTCAAAAGCAAAATAGTAATAATC 5766
DB 129 ATTTATACCCATTTAGGTTGGGTGGAATTTGTTCTCAAAAGCAAAATAGTAATAATC 188

QY 5767 TGGTATCTGCCCTATAACTCACAGTTGTATAAGAAAGTAGGCAGCACTCCTAGCATTTATAT 5826
DB 189 TGGTATCTGCCCTATAACTCACAGTTGTATAAGAAAGTAGGCCTTTCTCCTAGCATTTATAT 248

QY 5827 ATGATTTGGGTTTCTGAGTAACTGGGGAGTGTGTAGCTTTTGTAGTCTTGTAGTCTTGTAGTCT 5886
DB 249 ATGATTTGGGTTTCTGAGTAACTGGGGAGTGTGTAGCTTTTGTAGTCTTGTAGTCTTGTAGTCT 308

QY 5887 TAATTAGAA-AGTCTGTGGCTTTTACAGGCAATAGTCCCTTTGTCTGTTTGTCCCATGA 5945
DB 309 ATTAGAAAAGAAATATTGGCTTTTACAGGCAATAGTCCCTTTGTCTGTTTGTCCCATGA 368

QY 5946 TGCCCTTAAGTCTTTGGAGTCTCATTTAGAA-TTCTTTTCTCGAAGCATGCAAGTGT 6004
DB 369 TGCCCTTAAGTCTTTGGAGTCTCATTTAGAA-TTCTTTTCTCGAAGCATGCAAGTGT 428

QY 6005 ATCGCAATCTTACATGCTCCTCGTTTACCTGGCTTGTGTGCTGGTGTATTATTAAT 6064
DB 429 -----AATCAGTACTTGTCTCATTTATTTGTCTGTTTGTAGTCTGCTACTATTAT 483

QY 6065 GCATTTTCAGCATCATGCTTCTCTTACAAATATGATATTTTATTGTTACACTAAG 6124
DB 484 TATCTTTCAGGCTTTTCTCTTACAAATATGATATTTTATTGTTAGTGTAAAGCTAAG 543

QY 6125 GTGTTGA-TCATGTATCTGCTCCCTGTAAGAAATTAATTAATTAATTTTCCAGA 6175
DB 544 GCATTTGATCATGTATCTGCTCTTATTAATGAATTAATTAATTAATTTTCCAGA 595

RESULT 9
ABX71437
ID ABX71437 standard; cDNA; 2976 BP.
XX
AC ABX71437;
XX
DT 06-MAR-2003 (first entry)
XX
XX Human cDNA encoding a novel tyrosine phosphatase, NHP6.
DE Human; ss; gene; chromosome 15; chromosome 3; tyrosine phosphatase;
KW Ig super family protein; gene therapy; NHP; novel human protein.
XX Homo sapiens.
XX
XX US6465632-B1.
XX

PD 15-OCT-2002.
XX 08-JUN-2001; 2001US-0877730.
XX 09-JUN-2000; 2000US-210607P.
XX (LEXI-) LEXICON GENETICS INC.
XX Walke DW, Scoville J, Turner CA, Friedrich G, Abuin A;
PI Zambrowicz B, Sands AT;
PI WPI; 2003-147071/14.
DR P-PSDB; ABU54205.
XX
PT Novel isolated nucleic acid which encodes a novel human protein that
PT shares sequence similarity with animal phosphatases, that is useful for
PT generating antibodies, and as reagents in diagnostic assays -
XX
PS Disclosure; Column 49-52; 58pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule comprising a
CC nucleotide sequence encoding a novel human protein (NHP) that shares
CC sequence similarity with animal phosphatases (in particular tyrosine
CC phosphatases and is a member of the Ig superfamily). The disclosed NHP
CC polynucleotide sequences (from genes located on either chromosome 15
CC or chromosome 3) are useful for identifying coding sequences and in the
CC identification of biologically relevant splice junctions. The NHP
CC polynucleotide sequences are useful in gene therapy, for detecting
CC mutant NHPs or inappropriately expressed NHPs, for the diagnosis of
CC disease, for screening drugs effective in treatment of symptomatic or
CC phenotypic manifestations of perturbing the normal function of NH.
CC Complementary sequences of the NHP polynucleotides used in conjunction
CC with PCR to screen libraries, isolate clones and prepare cloning and
CC sequencing templates. Labeled NHP nucleotide probes can be used to screen
CC a human genomic library which is helpful for identifying polymorphisms,
CC determining the genomic structure of a given locus/allele and designing
CC diagnostic tests. The probe sequences also have use in defining and
CC monitoring both drug action and toxicity. Oligonucleotides
CC complementary to NHPs may encode or act as NHP antisense molecules, or
CC may be used as part of ribozyme and/or triple helix sequences.
CC The present sequence encodes an NHP.
XX
SQ Sequence 2976 BP; 844 A; 697 C; 653 G; 781 T; 1 other;

Query Match
Best Local Similarity 5.1%; Score 316.6; DB 25; Length 2976;
Matches 1210; Conservative 47.4%; Pred. No. 3.1e-71;
Matches 1210; Mismatches 1270; Indels 72; Gaps 6;

QY 300 AGACGCGATGATGAGGAGCTCTTATGATCTGGAAGTCTACTGAGGCGAGCTATTCCTG 359
DB 45 ATACATCAGTGGTGGAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 104
QY 360 TCTGGCCACAGCCCGCTAGAGTGGTGGCCAGCCAGGTTGCTGTGTTCAAGCTTGGCAC 419
DB 105 CTTGGCAATGAACAAATATGAGGCAATCTTAGTCAAAAAGCTCATCTTGGCTTATCAAC 164
QY 420 ACTCGAAGACTTCTCTGCAACCCGAGTCCAGATTCTGGAGGAGGAGGAGGAGGAGGAGG 479
DB 165 TATTTCTGATTTGAGTCCAGCCAAATTTCCAGTGGTCCAGGAGTGGAGTGGCTGCG 224
QY 480 CTTTGAATGCCACCAAGGCGCTTCCAGCCGCCCATCTATTCTGGGAAAGAGCACAGGT 539
DB 225 ATTTGATGCAAGATTTTCATCCACCCCTCTGCACTATACATGAGGAGTTCATATGGAC 284
QY 540 GACCGTGGCTGAG---GAGCCCGGCTCATCATCTTCCCAAGTGGCTCTCCAGATCCT 596
DB 285 AACTCTACCTATGATGAGCAGGATAAATCTGCCCTACCAACAGGAGTATTGCAGATCTA 344
QY 597 AGATGTCACGACAGTGTGAGGCTCTAGCGTGGTGGCCACCAATTCAGCCGCCA 656
DB 345 TGATGTACGCCAAAGGAGTCTGGAAATATCTGTTGTTATGCTGCACTGTAGCCACCG 404
QY 657 ACGATTACGCCAGGAGGCTCGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTACCG 716

DB 405 ACGTAAAGTATGAGGCGCTCGCTAACTGTGATTCCAG-----CTAAGGAGTCAAAATC 458
QY 717 GGGGCGAGGATGTGGTCAATTGGGCGAGCCCGACAGAGAACACACCGGTAGTGTCTGGACAGAA 776
DB 459 CTTCCACACACCACCAATATTAGCGGTCCACAGAACATAACAATCTCTTCATCAGAC 518
QY 777 TGTAGTATGAGTGGCTCTGTGAGCCCTCTGTGAGCCCAACCCCTTTTGTGTCTGGGTCCGACA 836
DB 519 TGTAGTTTGGAAATGCAATGGCCACAGGAATCCCAACCAATCATTTCTGGAGCGGCT 578
QY 837 GGATGGAAGCCTAT---CTCCACGGATGTCTGTCTGGGCGGACCAATCTACTCAT 893
DB 579 TGATCACAAATCCATTCATGTCTTAACTACTCGGTACTTGGAAATGGTAATCTCATGAT 638
QY 894 CGCCAGCGCGAGCTCGGCACTCTGGAGTCTATGTCTGCGAGCGAGCAACAGCCCTCCAC 953
DB 639 ATCTGATGTAGGCTACAAACATGCTGGAGTATATGTTTGTGCGGCCACTACCCCTGGGAC 698
QY 954 GCGTGAATTCGCCACTCGGCTCGGCTCGGAGTCTCGGAGTGTCTGTGCCCCAGCCATCTCGCA 1013
DB 699 ACGCAACTTTACAGTTGCTATGCGCAACTTTAACTGATTAAGTCTCTCTCTTCAATTGTTGA 758
QY 1014 GGCAACCGAGCGCTCTCGGAGCGCGGCGACGCGCGCTTGTGTGCGGCGCTC 1073
DB 759 ATGGCCAGAAAGTTTAAACAGGCGCTCGAGCTGCGACTGCTCGATTTGTGTGTCAGGACA 818
QY 1074 CGGCGAGCCACGCGCGCTGCACTGCTGCAAGCGGATCCGTTTGGACCCAAATGG 1133
DB 819 AGGAATCCCTCTCCCAAGATGTCTGTTGAAATGGAAGAGATACATTCGAATGG 878
QY 1134 GCGCGTCAAGGTGCGAGCGCGCTGCGGAGCTTGGTCACTCATCTCAGATCGGCTCGCAGA 1193
DB 879 TAGAATTAATAATGTACAACAGT-----AAATTGGTAATTAACAGATTTATCTCGAAGA 932
QY 1194 CGCTGGCTATCAACAGTGGTGTAGCAGAAAAACAGCGGGAATCTGCTGTGCGCTGCGCC 1253
DB 933 TGATGCTATTTATCAGTGCATGCGTGAAGATGAGCAAGGATCTATTTATCTAGAGCCAG 992
QY 1254 CTTGCGGTAGTGTGCGGAGGCGCTGCCAGCGCCCGAGCTCGGCTCAGAGCCACGCC 1313
DB 993 ACTGACTGTGTGATGTCAGAAAGACAGACCAGTGTCTCCCTTAAATGTACATGTGAAC 1052
QY 1314 GCTGAGCAGACTCTCTGTGCTGTGGCTGCGGAGCGGCTGAGTTGTCACAGCGAGCAAT 1373
DB 1053 CATGTCAGCTCAGCCATCTTTTAGCTGCGAGAGGCCACTTTATTAATTCAGACAAAGT 1112
QY 1374 CATTTGGCTTCTCTTCACTAACAAAGGCAAGGGAGTGGACAAATGTGGAGTACCAGTT 1433
DB 1113 CATTTGCCCTATTTCTGTACACTACATGAAAGCAGAAAGTTTAAATTAATGAAGATATCAAGT 1172
QY 1434 TGCAATAAACAAATGACACCAACAGAGCTGCGAGTTTGGGACCTGGAAACCAACACAGGATTA 1493
DB 1173 AGTCATCGGAAATGACCAACTCATTTATTTATTTGATGACTTAGAGCCTGCGCAGCAATTA 1232
QY 1494 TGAGTTTCTAGCTGTGGCTTACTCCAGCTGCGGCGCAGCGAGCACTCTCAGCCAGCCCT 1553
DB 1233 TACTTTCTACATTTAGCATATATGCCAATGGAGCGCCAGCAGATGTCTGACCATGTGAC 1292
QY 1554 GGTGCATACATGGAACGATGTCGCCAGCGAGCAGCCCGAGCTTACTTTGTCTGAGCCCAA 1613
DB 1293 ACAGAACTCTAGAGGATGTTCCCTGAGACCTCTCGAAATTTAGTTTGAAGAATCGAAG 1352
QY 1614 CCCCTCGGACATCAGGTTGGCATGTGCTGCCCTGCCCTTCCAGCTTCCAGCAATGAGCAGGT 1673
DB 1353 TCCCACTGATATTTCTCTCTGCTGCTGCAATCCAGCCAAATATCGCGGGGCGCAAGT 1412
QY 1674 GCTCAAGTACAAAGTAGAGTACGGTTTGGGGAAGGAGATCAGGTTTCTCCAACCGAGGT 1733
DB 1413 GGTGCTGTATGCTGTGCTTTCCGCTTAAAGTACTGAGAAATTCAAATCCAAGTCTTGAGCT 1472
QY 1734 GCCTGGAAATGAGACACAACTTAGCTTAAATCTCACTTCCAGCCAAACAAAGTGTACCGAGT 1793

Db 1473 CCGGGGACCAACGATGAGTACCTTTTGAAGGCTGAAACCTGACAGTGTCTACCTGGT 1532
Qy 1794 CCGGATTTACAGTGGCACTGGCTGCTATGAGTCCCTTCTCAGTGGATGACACAG 1853
Db 1533 TCGGATTTACTGCTGCCACAGAGTGGGCTGGGAGATCATCATGATGGATTTCAATG 1592
Qy 1854 GACACCTGGTGTGCACAAACAGAGCCATGTTTCCCTTTGCCCTGCGAGAAATTGAAGGTGAG 1913
Db 1593 GACGCCAAAGCTACAAGCGTGAAGCCCTTAAGTCTCCAGAGTTGCATTTGGAGCTCT 1652
Qy 1914 GGCAAAGATGGAGTCCCTGCTGTGTGTATGCGAGCCGCCCTTCAACCCACCCAGATCTC 1973
Db 1653 GAACTGTACCAACCAATTTCTGTGAGGTGGCAGCAAGATGTAGAGGACACAGCTCTATTCA 1712
Qy 1974 TGGATACAACTCTACTGGGAGAGGTGGGAAACAGAGGAGGAGGACAGATGTGTACCGCCC 2033
Db 1713 GGGCTACAGCTGTACTACAGGAAGAGGGGACGAGGAGAA----- 1754
Qy 2034 CCCAGGGGGTCTGGAGATCAAGCTTGGGACGTGGGCCCGTGGCTGGAAGAAAGT 2093
Db 1755 -----TGGGCCCAATTTCTTGGATACCAAGGA 1781
Qy 2094 GAAGCAGTATGAACCTGACCTGATGTTTCTTGGAGCCGCTACAGAGTGAAGTCTGTAGC 2153
Db 1782 CCTACTCTATACTCTCAGTGGCTTAGACCCGAGAAATAATCATGTGAGACTCTCTGGC 1841
Qy 2154 TTTCAACAAACAGAGACCGCTAGCTGTGTGTGTGGAAGGCGAGAGGAGGCGCC 2213
Db 1842 TTACAACAAATAGACATGGCTATCAGGAGATCAGACTGTGAGCAC-----TCCAGGATG 1898
Qy 2214 CAGCCAGACTGCTCTATCAGAGGGGCGCACCGCTGCTCTGCCCCATGTCACAGCAGA 2273
Db 1899 CGTGTCTGTGTGATGCGATGTGCTCTCTCACACACACCCACCATCTCTATGGGAA 1958
Qy 2274 GTCAACAGCTTCACTTCCATTTGGCTTGGTGGAGAAAGCCAGACTTTTACCACTGTCAA 2333
Db 1959 GGGTAAACCTCTCTTCTTCTTCTGCTGCTGAGGAGGCGCTGCACTTACCGCTGCACA 2018
Qy 2334 GATTTGCTACACTGACCTGCTGCGCCCTGGGCGCTCAGGATGCTTCCCTGGTCCAC 2393
Db 2019 AATCATTAACCTACACCACTCGCTGTATCTCTTGGCTGCGAGATGCTTCTTGGTCT 2078
Qy 2394 CTACTATACAGCTCTCGAGAGACATTTCTATTTGGCGGCTGAAACCAATTTACCAAGTA 2453
Db 2079 GTACTCTCAACATCAGAACTCATATGTTGGTTCAGGTCTAGAACCAACCAACAAATA 2138
Qy 2454 CGAGTTTCCGCTACAGTCCCAACGAGTGGATATGAGTGGGCGCTTGGCTCCGTCGPAGA 2513
Db 2139 CGAATTTGCGCTTCGATACATGTGGATCAGCTTTCCAGTCTTGGAGCCCTGTAGTCTA 2198
Qy 2514 ACGTCCACCTGCTGACCGGCTTCAACACTCTCTTGAACCTGCGCTGAGCCCTT 2573
Db 2199 CCATTCTACTCTTCCAGAGCACCAGAGGCGCCACCAAGTTGGAGTAAAGTGAATTAAT 2258
Qy 2574 GACACCATCAACCGTGGTTTACACTGTGTGTCGCCCCACAGGCGCCAAATGGTGAATGT 2633
Db 2259 AGAGATGACACTGCCCTGGTTTCTTGGAAACCCCTGTATGGCCAGAAACAGTTGTGAC 2318
Qy 2634 GGAGTATCTAATTTCTTACAGCAACCAACACCCAGCCGGAACACCACTGGACATGCT 2693
Db 2319 CCGCTATACTATCTTATATGATCATAGAGGCGCTGGAATGTCAGGAGTGGCAGGCTT 2378
Qy 2694 CACACAGAGGGAACATCTTCTAGTGAGAGGTCCATGGCTAGAGATGACATCGGTA 2753
Db 2379 ACACCGTGAAGGGCAATAACCATGGCTTTGCTAGAAAACTTGTGTAGCAGGAATGTGTA 2438
Qy 2754 TTTCTCAAGATGGAGCCGACAGAGGTGGGCGCTTGGCGCTTTTCCGCTTGCAGGA 2813
Db 2439 CATTTGTCAGATATCTGATCCATGAGTGGGAGAGGACCCCTTTTCAATTTCTGTGGA 2498
Qy 2814 TGTGATTACTTGAAGAGACATTTCTCAGATCTC 2846
Db 2499 -GCTGGCAGTACTTCCNAGGAAACCTCTGATC 2531

RESULT 10

ABX71432
ID ABX71432 standard; cDNA; 3210 BP.

XX AC ABX71432;

XX DT 06-MAR-2003 (first entry)

XX DE Human cDNA encoding a novel tyrosine phosphatase, NHP1.

XX KW Human; ss; gene; chromosome 15; chromosome 3; tyrosine phosphatase;
KW Ig super family protein; gene therapy; NHP; novel human protein;
KW SNP; single nucleotide polymorphism.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT variation 76

FT FT /tag= a /standard_name= "single nucleotide polymorphism"

FT FT 706

FT FT /tag= b /standard_name= "single nucleotide polymorphism"

XX US6465632-B1.

XX PD 15-OCT-2002.

XX XX 08-JUN-2001; 2001US-0877730.

XX XX 09-JUN-2000; 2000US-210607P.

XX XX (LEXI-) LEXICON GENETICS INC.

XX XX Walke DW, Scoville J, Turner CA, Friedrich G, Abuin A;

PI Zambrowicz B, Sands AT;

XX WPI; 2003-147071/14.

DR P-PSDB; ABU54200.

XX Novel isolated nucleic acid which encodes a novel human protein that
shares sequence similarity with animal phosphatases, that is useful for
generating antibodies, and as reagents in diagnostic assays -
Disclosure; Column 13-18; 58pp; English.XX The invention relates to an isolated nucleic acid molecule comprising a
nucleotide sequence encoding a novel human protein (NHP) that shares
sequence similarity with animal phosphatases (in particular tyrosine
phosphatases and is a member of the Ig superfamily). The disclosed NHP
polynucleotide sequences (from genes located on either chromosome 15
or chromosome 3) are useful for identifying coding sequences and in the
identification of biologically relevant splice junctions. The NHP
polynucleotide sequences are useful in gene therapy, for detecting of
mutant NHPs or inappropriately expressed NHPs, for the diagnosis of
disease, for screening drugs effective in treatment of symptomatic or
phenotypic manifestations of perturbing the normal function of NH.
Complementary sequences of the NHP polynucleotides used in conjunction
with PCR to screen libraries, isolate clones and prepare cloning and
sequencing templates. Labeled NHP nucleotide probes can be used to screen
a human genomic library which is helpful for identifying polymorphisms,
determining the genomic structure of a given locus/allele and designing
diagnostic tests. The probe sequences also have use in defining and
monitoring both drug action and toxicity. Oligonucleotides
complementary to NHPs may encode or act as NHP antisense molecules, or
may be used as part of ribozyme and/or triple helix sequences.

XX The present sequence encodes an NHP.

SQ Sequence 3210 BP; 895 A; 761 C; 716 G; 836 T; 2 other;

Query Match

5.1%; Score 316.6; DB 25; Length 3210;

Qy	1074	CGGGAGCCACGGCCCGCGCTGCACTGGCTGCACGACGGGATCCGTTGGGACCCCAATGG	1133
Db	819	AGSAAATCCCTCTCTCCCAAGATGTATGGTTGAAATAATGGAAGGAGATACATTCGAATGG	878
Qy	1134	CGCGGTCAAGGTGACGAGCGGTGGCGGACGTTCGTATCACTCAGATCGGCTGCAGGA	1193
Db	879	TAGAAATTAAATGTACACAGT-----AAATTGGTAATTACACGAATATTCTCGAAGA	932
Qy	1194	CGCTGGCTATACACAGTGCCTGAGAGAAACAGCGCGGAACTGCCTGTGCGCTGCGCC	1253
Db	933	TGATGCTATTTATCAGTGCATGGCTGAGAAATAGCAAAGGATCTATTTATCTAGAGCCAG	992
Qy	1254	CTTGGCGGTAGTGTGGCGAGGGGTGCCCCAGCGCCCGGACTCGGGTCAACAGCCAGCC	1313
Db	993	ACTGACTGTGTAGTGTGACGAAGACAGACCCAGTGCCTCCCTATTAATGTATCATGCTGAAAC	1052
Qy	1314	GCTGAGCAGCTCTCTGTGCTGGTGGCTGGGAGCGGCTCGAGTTGCACAGCGGAGCAAT	1373
Db	1053	CATGTCAAGCTCAGGCATCTTTTAGCTGGGAGGGCCACTTTATTAATTCAGACAAAGT	1112
Qy	1374	CATTGGCTTCTCTCTTCATACCAAAAGCGAAGGGAGTGACAAATGTGGAGTACCAAGT	1433
Db	1113	CATTGGCTATTCTGTACACTACATGAAAGCAGAGGTTTAAATTAATGAAGTATCAAGT	1172
Qy	1434	TGCAGTAAACAAATGACACACAGAGCTGCAGTTTCGGGAACCTGGAAACCAACACGGATTA	1493
Db	1173	AGTCATCGGAAATGACACAACTCATATATATTATTGATGACTTAGAGCTGCGCAGCAATTA	1232
Qy	1494	TGAGTTCTACGTGTGGCTTACTCCAGCTGGGGGCGAGCGGAACCTCCAGGCCGAGCCCT	1553
Db	1233	TACTTTCTACATTTGTAGCATATATGCCAATGGGAGCCAGGCAGATGTCTGAACCATGTGAC	1292
Qy	1554	GGTGATACACTGCACAGATGCCAGCGCAGACCCCGAGCTTACCTTGTTCACGCCCAA	1613
Db	1293	ACAGATACTCTAGAGGATGTTCCCTGAGACCTCTTGAATTAAGTTTGAACAGTGAAG	1352
Qy	1614	CCCCTCGGACATCAGGGTGGCAATGGCTGCCCCCTCCAGCTGCAGCAATGACACAGGT	1673
Db	1353	TCCCACTGATATTCTCATCTCTGGCTGCCAATCCAGCCAAATATCGCGGGGCCAAGT	1412
Qy	1674	GCTGAAGTACAAGATAGAGTAGCGTTGGGGGAAGGAAGATCAGGTTTTTCCACCGAGGT	1733
Db	1413	GGTGCTGTATCGCTGTCTTTCCGCTAAAGTACTGTAGAAATTCAAATCAAGTTCTGGAGCT	1472
Qy	1734	GCCTGGAAATGAGACACAACTTACGTTAACTCACTTCAGCCAAACAAAGTGTACCGAGT	1793
Db	1473	CCCGGGACACCGATGATGATACCTTTTGGGAAGCCCTGAAACCTGACAGTGTCTACCTGGT	1532
Qy	1794	CCGGATTTTCAGCTGGCATGCGCTGGCTATGGAGTCCCTTCTCAGTGGATTCAGCAGACAG	1853
Db	1533	TCGGATTACTCTGCCACAGAGTGGGGCTGGGAGAGTCATCAGTATGGACTTCACATAG	1592
Qy	1854	GACACTGGTGTGCACACACAGAGCCATGTTCCCTTTGGCCCCCTGCAGAAATGAAGGTGAG	1913
Db	1593	GACGCCAAAGACTAAGACGCTGAAGGCCCTTAAGTCTCCAGAGTTGCATTTGGAGCCCTCT	1652
Qy	1914	GGCAAGATGGAGTCCCTGGTGGTGTATGGCAGCGCCCTCCCTCACCCACCCAGATCTC	1973
Db	1653	GAACTGTACCAACCATTTCTGTGAGGTGGCAGCAAGATGTGAGGACACAGCTGCTATTCA	1712
Qy	1974	TGGATACAAACTCTACTTGGGAGAGGTGGGAAACAGAGGAGGAGGCAGATGGTGAACGCC	2033
Db	1713	GGGCTACAAGCTGTACTACAAGGAAGAGGGCAGCAGAGNA-----	1754
Qy	2034	CCCAGGGGTCTGTGGAGATCAAGCTTGGGACGTGGGCGCGCTGGGCTGAAGAGAAAGT	2093
Db	1755	-----TGGGCCCATTTTCTTGGATACCAAGGA	1781
Qy	2094	GAAGCAGTATGAACTGACCCAGTTAGTCCCTGGCAGGCGGTACAGAGTGAAGCTCGTAGC	2153
Db	1782	CCTACTCTATACTCTCATGTGGCTTAGACCCCGAGGAAGAAATATCATGTGAGACTCTCGGC	1841

Qy	2154	TTTCAAAACACGAGGAGCGCTACGCTCTGTGTGAGAGGCGAAGCGAGGAAGCGCC	2213
Db	1842	TTTCAAAACACATAGACGATGGCTATACGGCAGATCAGACTGTCAGCAC---	1998
Qy	2214	CAGCCAGACCTGCCTATCCAGAGGGGGCCACCGCTGCCTCTCGCCCATGTCCACGCAGA	2273
Db	1899	CGTGTCTGTTGCGTATCGCATGTGCTCTCCACACACCCACCATCTCTATGCGAA	1958
Qy	2274	GTCAAAACAGCTCCACATTCGATTTGGCTTCGGTGGAGAGAGCCAGACTTTACCACTGTCAA	2333
Db	1959	GGCTAACACCTCATCTTCCATCTCTCTGCACTCGAGGAGGCGCTGCATTCACCGCTGCACA	2018
Qy	2334	GATTGTCAACTACACTGTACGCTTCGGGCCCTCGGGGCTCAGCAATGCTTCCCTCGTCCAC	2393
Db	2019	AATCATTAACACACCATCGCTGTAATCTCTGTTGGCTCGCAGAAATGCTTCTTTGGTTCT	2078
Qy	2394	CTACTATACAGCTCTGGAGAGACATTCATCTGGCGGCTGAAACCAATTTACCAAGTA	2453
Db	2079	GTACCTTCAAACATACGAACTCACATGTTGGTTCAAGGCTTAGAAACCAACACCAATA	2138
Qy	2454	CGAGTTGGCGTACAGTCCACCGAGTGGATATGGATGGGCCCTTTGGCTCCGTCGTAGA	2513
Db	2139	CGAATTGGCGTTCGATTACATGTGGATCAGCTTCCAGTCTTTGGAGCCCTGTAGTCTTA	2198
Qy	2514	ACGCTCCACCTCGCTGACCGGCTTCAACACACTCTTCTTGACCTCGGCTCGAGCCCCCT	2573
Db	2199	CCATTCTACTCTTCCAGAGCACCGCAGCGCCCAACAGTTGGAGTAAAGTGACATTAAT	2258
Qy	2574	GACACCATCACCGTTCGGTTACACTGGTGTCCCCCAACGAGGCCCAATGGTGAGATTGT	2633
Db	2259	AGAGATGACACTGCGCTGGTTCTTGAAACCCCTTGATGGCCCAAAACAGTTGTGAC	2318
Qy	2634	GGAGTATCTAATTCTCTACAGCAACACCAACACCCAGCCCGAACACACAGTGGACACTGCT	2693
Db	2319	CCGCTATFACTATCTTATATGCATCTAGGAAGGCTCGAATTGACGAGAGTGGCAGGTCTT	2378
Qy	2694	CACCAAGAGGGAAAATCTTTAGTGCAGAGGTCCATGGCCTAGAGAGTGACACTCGGTA	2753
Db	2379	ACACCGTGAAGGGGCAATAACCATGGCTTTGCTAGAAAACCTTGGTAGCAGGAAATGTGA	2438
Qy	2754	TTTCTTCAAGATGGGAGCCGCAACAGAGGTGGGGCCCTGGGCCCTTTTCCCGCTTGAGGA	2813
Db	2439	CATTGTCAAGATATCTGCATCCAAATGAGGTGGGAGGAGNCCCTTTTCAAAATCTGTGGA	2498
Qy	2814	TGTGATTACTCTGCAAGAGACATTTCCAGACTC	2846
Db	2499	GCTGGCAGTACTTCCAAAGGAAACCTCTGAATC	2531

RESULT 12
ARX71447

ABX/144/
ID ABX71447 standard: CDNA: 3874 BP.

AC ABX71447;

DT 06-MAR-2003 (first entry)

DE Human cDNA encoding a novel tyrosine phosphatase.

KW Human; ss; gene; chromosome 15; chromosome 3; tyrosine kinase

XX is super family protein; gene therapy; nnt; novel human protein.
OS Homo sapiens.

PN US6465632-B1.

15-OCT-2002.

08-JUN-2001;

AA 09-JUN-2000; 2000US-210607P.

PA (LEXI-) LEXICON GENETICS INC

XX Walke DW, Scoville J, Turner CA, Friedrich G, Abuin A;
PI Zambrowicz B, Sands AT;
XX WPI; 2003-147071/14.
XX Novel isolated nucleic acid which encodes a novel human protein that
PT shares sequence similarity with animal phosphatases, that is useful for
PT generating antibodies, and as reagents in diagnostic assays -
XX Disclosure; Column 107-112; 58pp; English.
XX The invention relates to an isolated nucleic acid molecule comprising a
CC nucleotide sequence encoding a novel human protein (NHP) that shares
CC sequence similarity with animal phosphatases (in particular tyrosine
CC phosphatases and is a member of the Ig superfamily). The disclosed NHP
CC polynucleotide sequences (from genes located on either chromosome 15
CC or chromosome 3) are useful for identifying coding sequences and in the
CC identification of biologically relevant splice junctions. The NHP
CC polynucleotide sequences are useful in gene therapy, for detecting of
CC mutant NHPs or inappropriately expressed NHPs, for the diagnosis of
CC disease, for screening drugs effective in treatment of symptomatic or
CC phenotypic manifestations of perturbing the normal function of NHP.
CC Complementary sequences of the NHP polynucleotides used in conjunction
CC with PCR to screen libraries, isolate clones and prepare cloning and
CC sequencing templates. Labeled NHP nucleotide probes can be used to screen
CC a human genomic library which is helpful for identifying polymorphisms,
CC determining the genomic structure of a given locus/allele and designing
CC diagnostic tests. The probe sequences also have use in defining and
CC monitoring both drug action and toxicity. Oligonucleotides
CC complementary to NHPs may encode or act as NHP antisense molecules, or
CC may be used as part of ribozyme and/or triple helix sequences.
CC The present sequence encodes an NHP.
XX
SQ Sequence 3874 BP; 1077 A; 938 C; 897 G; 960 T; 2 other;
Query Match 5.1%; Score 316.6; DB 25; Length 3874;
Best Local Similarity 47.4%; Pred. No. 3.6e-71;
Matches 1210; Conservative 1; Mismatches 1270; Indels 72; Gaps 6;
QY 300 AGACAGGATGATGAGAAAGCTCTTAGGATCTGGAGGTCACTGAGGGGAGCTATTCCTG 359
DB ATACATCAGTGGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 528
QY 360 TCTGGCCCAAGCCGCTAGGAGTGGTGGCCAGCCAGGTTGCTGTTCAAGTTCGCCAC 419
DB CTTGGCAATGAACAAATATGGAGCCATCTTAGTCAAAAAGCTCATCTTGCCTTATCAAC 588
QY 420 ACTCGAAGACTTCTCTGACCCCGAGTCCAGATTTGTGGAGGAGAAACGGGACAGCAG 479
DB TATTTCTGCATTTGAACTCCAGCCAAATTTCCACTGAGGTCCACGAAAGGTGAGTTGCTG 648
QY 480 CTTTGAATGCCACACCAAGGGCTTTCAGGCCCATCATTTACTTTGGGAAAGGACAGGT 539
DB ATTTGATGACAGATTTTCATCCCACTCCAGTCAATCAATGAGGAGTTTCAATCGGAC 708
QY 540 GACGTCCTGAG---GAGCCCGGCTCATCTTCCCAAGTGGCTCTCCAGATCCT 596
DB AACTCTACCTATGACTATGACAGGATACTGCCCTTACCACAGGAGTATTCAGATCTA 768
QY 597 AGATGTCACGAGATGATGACAGGCTCTTACCGCTGGTGGCCACCAATTCAGCCCGCCA 656
DB TGATGTCAGCCAAAGGATTTCTGMAATATCTGTTATGCTGCCACTGTAGCCCAACCG 828
QY 657 ACGATTGACGAGGAGGCTCGCTCACTGTGGCCCTCAGAGGGTCTTTTGAGGAGTACACG 716
DB ACGTAAAGTATGAGGAGGCTCGCTCACTGTGATTTCCAG-----CTAAGGAGTCAAAATC 882
QY 717 GGGGACAGATGTTGTTGGAGCCCGCCAGAGAACACACCGTAGTGTCTGGACAGAA 776
DB CTTCCACACACCAACCAATTTAGCAGGTCCACAGAACATTAACCAATCTCTTCATCAGAC 942
QY 777 TGTAGTGTGAGTGGCTCTGCTGACCCCGCCACCCCTTTTGTGTCTGGGTCCGACA 836

DB 943 TGTAGTTTGGAAATGCAATGGCCACAGGAAATCCCAAAACCAATCATTTCTTGGAGCGCCT 1002
QY 837 GGATCGAAAGCCTAT---CCACCGGATGTCATCTCTGGGCGGACCAATCTACTCAT 893
DB TGTATCAAAATCCATTTGATGTTTAAATCTCGGGTACTTGGAAATGGTAATCTCATGAT 1062
QY 894 CGCCAGCGCGCAGCCTCGGCACCTCTGGAGTCTATGTCTGCCGAGCAACAAAGCCCTCAC 953
DB 1063 ATCTGATGTCAGGCTACAAACATGCTGGAGTATATGTTGTGGGCGCACCTACCCCTGGCAC 1122
QY 954 GGTGATCTTGGCACTGGCGGCTGAGCTCGAGTCTCGAGTCTGTGTCGCCAGCCATCTCGCA 1013
DB 1123 AGCCAACTTTACAGTTGCTATGGCAACTTAACTGTAATAGTCTCTCTTCAATTTGTTGA 1182
QY 1014 GGCACCCGAGGCGCTCTCGGAGCGGGGCCAGCACCGCGCTTCTGTGTGCGGGCGTCTC 1073
DB 1183 ATGGCCAGAAAGTTTAAACAGGCTCTGAGCTGGCACTGCTCGATTTGTGTGTCAGGAGA 1242
QY 1074 CGGGGAGCCACCGGCCCGCTGCTGCTGCTGCAACCGGATCCCGTTTGGCAACCAATGG 1133
DB 1243 AGGAATCCCTCTCCAAAGATGTCATGTTGTTGAAAATGGAAGAAATATATTCGAATGG 1302
QY 1134 GCGCTCAAGGTGCAAGGGGGTGGCGGAGCTTGGTCTCATCTCAGATCGGCTCGCAGGA 1193
DB 1303 TAGAATTTAAATGTACAAAGT-----AAATTTGGTAAATTAACCAAGATTTATTCCTGAAGA 1356
QY 1194 CGCTGGCTACTACCACTGCTAGCAGAAACAGCGCGGGAACCTGCTGTGCGCTGCGCC 1253
DB 1357 TGATGCTATTTATCAGTGTGATGGCTGAGAAATAGCAAGGATCTATTTATCTAGAGCCAG 1416
QY 1254 CTTGCGGTAGTGGTGGCGAGGGGCTGCCAGCGCCCGCACTCGGGTCAACAGCCACGCG 1313
DB 1417 ACTGACTGTGATGATGTCAGAGACAGACCCAGTCTCCCTATATGATGATGCTGGAAC 1476
QY 1314 GCTGAGAGCTCCTCTGTGCTGGTGGCTGGAGCGGCTGATGTTGCAAGGAGCAAT 1373
DB 1477 CATGTCAAGCTCAGCCATCTTTTATAGCTGGGAGAGGCCACTTTATTAATTCAGACAAAGT 1536
QY 1374 CATTCGCTCTCTCTTCACTACCAAAAGCAAGGGGAGTGGCAATGTGGAGTACCAGTT 1433
DB 1537 CATTCGCTATTTCTGTACACTACATGAAGCAGAGGTTTAAATATGAAGAGTATCAGT 1596
QY 1434 TGCAGTAAACAAATGACACCAAGAGCTGCGAGGTTTGGGACCTGGGAACCCAAACAGGATTA 1493
DB 1597 AGTCATCGGAATGACAACTCATTTATTTATGATGATTTAGAGCTTGCAGCAATTA 1656
QY 1494 TGAGTTCTAGTGGTGGCTTCTCCAGCTGGGGGCGAGCGAACTCCAGGCCAGCCCT 1553
DB 1657 TACTTTCTACATTTGTAGCATATATGCCAATGGGAGCCAGAGATGCTGACCATGTGAC 1716
QY 1554 GGTGATACATCGGAGATGTCGCCAGCGCAGCACCCAGCTTACCTTGTCCAGCCCA 1613
DB 1717 ACAGAAATCTCTAGAGGATGTTCCCTCTGAGACCTCTCTGAAATTTAGTTTGAAGTCCAG 1776
QY 1614 CCGCTCGACATCAGGGTGGCATGGCTGCCCTTCCCTCGAGCTCCAGCTCAGCAATGGAAGT 1673
DB 1777 TCCCACTGATATCTCATCTCTGCTGCTCAATCCAGCCAAATATCGGGGGGCCAAGT 1836
QY 1674 GCTGAAGTACAAGATAGAGTACGTTTGGGGAAGGAAGATCAGGTTTTTCTCCACCGAGT 1733
DB 1837 GGTGCTGATCGCTTGTCTTTCCGCTAAAGTACTGAGAAATTCAAATCCAAAGTCTCGAGCT 1896
QY 1734 GCTGGAATGACACAACTTACGTTAACTCATCTTCAGCCGCAACAAAGTGTACCGAGT 1793
DB 1897 CCGGGGACCAACGCAATGAGTACCTTTGGAAGGCCCTGAAACCTGACAGTGTCTACCTGGT 1956
QY 1794 CCGGATTTTCAGCTGGCAGCTGGCTGCTATGGAGTCCCTTCTCAGTGTGATGCAAGCAG 1853
DB 1957 TCGATTTACTGTGCCACAGAGTGGGGCTGGAGAGTCTATCAGTATGAGCTTCACATAG 2016
QY 1854 GACACCTGTTGTCACAAACAGAGCCATGTTCCCTTTGCCCCCTGCAAGATTTGAAGGTGAG 1913

Db 2017 GACGCCAAAGCTACAGCGGTGAAGCCCTTAAGTCTCCAGAGTTGCAATTTGGAGCTCT 2076
Qy 1914 GGCAGAGTGAAGTCCCTGGTGGTGCATGGCAGCGCCCTCACCACCCACCCAGATCTC 1973
Db 2077 GAACCTGTACCACCTTTCTGTGAGGTGCAGCAAGATGTAGAGCACACAGCTGTATTCA 2136
Qy 1974 TGGATACAACTCTACTGGGAGAGGTGGGAACAGAGAGGAGGAGAGTGTGACCGCC 2033
Db 2137 GGGCTACAAAGCTGTACTACAAGGAAGAGGCGCAGCAGAGAA----- 2178
Qy 2034 CCCAGGGGCTGTGGAGATCAAGCTTGGGACGTGGGCGCTGGCGCTGAAGAGAAAGT 2093
Db 2179 -----TGGGCCCAATTTCTTGGATACCAAGA 2205
Qy 2094 GAAGCAGTATGAATGACCCAGTTAGTCCCTGGCAGGCGGTACGAGGTGAAGCTCTGTAGC 2153
Db 2206 CCTACTCTATCTCTCAGTGGCTTAGACCCCAAGAAATATCATGTGAGACTCTCTGGC 2265
Qy 2154 TTTCAACAAACACGAGGACGGCTAGCTGTGTGTGGGAAGGGCAAGACGAGAGAGCGCC 2213
Db 2266 TTAACAACAACATAGACGATGGCTATCAGGCGAGATCAGACTGTGAGCAC---TCCAGGATG 2322
Qy 2214 CAGCCAGACTGCTATCCAGAGGGGGCGACCGCTGCTCTGCCCATGTCCAGCGAGA 2273
Db 2323 CGTGTCTGTTGATGATGGTCCCTCTCCACACCCACCCACCATCTCTATGGAA 2382
Qy 2274 GTCAACAGCTCCACTTCATTTGGCTTTCGGTGGGAAGAGCAGACTTTTACCACTGTCAA 2333
Db 2383 GGCTAACACCTTCATCTTCCTGCTACTGGAGGAGGCTGTGATTCACCGCTGCACA 2442
Qy 2334 GATTGTCAACTACACTGTACCTGGCCCTTGGGGCTCAGGAATGCTTCCCTGGTCA 2393
Db 2443 AATCATTAATACACCATCTGCTGTATCTGTTGGCTGCGAATGCTCTTTGGTTCT 2502
Qy 2394 CTACTATACAGCTCTGGAAGACATCTCATTTGGGGCTGGAACCATTTTACCAAGTA 2453
Db 2503 GTACCTTCAACATCAGAACTCACTATGTTGTTCAAGGTTAGAACCAACACCAATA 2562
Qy 2454 CGAGTTTGGGTACAGTCCACGAGTGGATATGATGGGCGCTTTGGCTCGTCTAGA 2513
Db 2563 CGAATTTGGCTTGAATACATATGATGATCAGCTTTCAGTCTTGGAGCCCTGTAGTCTA 2622
Qy 2514 ACGTCCACCTGCTGACCGGCTTCAACACCTCTTCTGACCTCGCCCTGAGCGCCCT 2573
Db 2623 CCATTTCTACTCTTCAGAGACACGACGAGCGCCACCATGTTGAGTAAAGTGACATTAAT 2682
Qy 2574 GACACCATCCAGTTCGGTTACTGTGTGTCCTCCCAAGGAGCCCAATGTGTGATGT 2633
Db 2683 AGAGGATGACACTGCTGCTGTTCTTGGAAACCCCTGATGGCCCAAGAAACAGTTGTGAC 2742
Qy 2634 GGAGTATCTAATTTCTACAGCAACACACACCCAGCCCGAACACAGTGGACACTGCT 2693
Db 2743 CCGCTATCTATCTTATATGATCTAGAAAGGCTTGGATTCGAGAGATGGCGGCTT 2802
Qy 2694 CACCACAGAGGAAACATCTTTCAGTGCAGAGGTCCATGGCTAGAGAGTGAACCTCGTA 2753
Db 2803 ACACCGTGAAGGGCAATAACATGGCTTCTAGAAACATTTGTGACGAGAAATGTGTA 2862
Qy 2754 TTTCTTCAAGTGGAGCCCGACAGAGTGGGCGCTGGGCGCTTTTCCCGCTTGCAGGA 2813
Db 2863 CATTTCAAGATATCTGATCCAAATGAGGTGGGAGAGGACCTTTTCAAAATCTGTGGA 2922
Qy 2814 TGTGATTAATCTGCAAGAGACATTTCTGACTC 2846
Db 2923 GCTGGCAGTACTTCCAAAGGAACCTCTGAAATC 2955

RESULT 13

ID AAD36777 standard; cDNA; 2700 BP.

XX AAD36777;

XX

DT 30-OCT-2002 (first entry)
XX Human immunoglobulin superfamily protein (IGSFP)-1 cDNA.
XX Human; immunoglobulin superfamily protein-1; IGSFP-1; asthma;
XX immune system disorder; acquired immune deficiency syndrome; AIDS;
XX atherosclerosis; neurological disorder; Alzheimer's disease;
XX Parkinson's disease; developmental disorder; renal tubular acidosis;
XX anemia; muscle disorder; cardiomyopathy; myocarditis; cancer;
XX cell proliferative disorder; arteriosclerosis; hepatitis; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 1..2382
XX /tag= a
XX /product= "IGSFP-1"
XX sig_peptide 1..105
XX /tag= b
XX mat_peptide 106..2379
XX /tag= c
XX /product= "Mature IGSFP-1"
XX MO200240671-A2.
XX 23-MAY-2002.
XX 15-NOV-2001; 2001WO-US44974.
XX 16-NOV-2000; 2000US-249645P.
XX (INCY-) INCYTE GENOMICS INC.
XX Baughn MR, Lu DAM, Yue H, Elliott VS, Thangavelu K, Raskumar J;
XX Lu Y, Lo TP, Gururajan R, Gandhi AR, Arvizu C, Yao MG,
XX WPI; 2002-519384/55.
XX P-PSDB; AAE14781.
XX Novel human immunoglobulin superfamily polypeptide, useful in
XX diagnosis, prevention or treatment of immune system, neurological,
XX developmental, muscle and cell proliferative disorders -
XX Claim 5; Page 116-117; 122pp; English.
XX The present sequence is human immunoglobulin superfamily protein
XX (IGSFP)-1 cDNA. The IGSFP polypeptide and polynucleotide are useful
XX for diagnosing, treating or preventing disorders associated with aberrant
XX expression of IGSFP e.g. immune system disorders (e.g. acquired immune
XX deficiency syndrome (AIDS), asthma, atherosclerosis, psoriasis, uveitis),
XX neurological disorders (e.g. Alzheimer's disease, Huntington's disease,
XX dementia, Parkinson's disease), developmental disorders (e.g. renal
XX tubular acidosis, epilepsy, anaemia), muscle disorders (e.g.
XX cardiomyopathy, myocarditis), or cell proliferative disorders (e.g.
XX arteriosclerosis, cirrhosis, hepatitis, and cancer). The polypeptide and
XX polynucleotide are also useful for assessing the effects of exogenous
XX compounds on their expression. The polypeptide is useful in drug
XX screening techniques, to analyse the proteome of a tissue or cell type,
XX as elements on a microarray. The polynucleotide is useful for creating
XX knock-in humanised animals or transgenic animals to model human diseases,
XX in somatic or germline gene therapy, to generate a transcript image of a
XX tissue or cell type, for detecting differences in the chromosomal
XX location due to translocation, inversion among normal, carrier or
XX affected individuals, and as hybridisation probes for mapping naturally
XX occurring genomic sequences.
XX Sequence 2700 BP; 528 A; 859 C; 815 G; 498 T; 0 other;

Query Match 5.1%; Score 316.4; DB 24; Length 2700;

Best Local Similarity 55.0%; Pred. No. 3.3e-71;

Matches 692; Conservative 0; Mismatches 551; Indels 15; Gaps 3;

Qy 342 TGAGGCGAGCTATTCTCTGCTGCGCCACAGCCCGCTAGGAGTGTGCGCAGCGATTGC 401

Db 333 TGAAGGTGATATGAGTGTGTGGCCAGAAACCGCTTTTGGGCTGTGTGGTCTGAGCCGGAAGGC 392
QY 402 TGTGGTCAAGCTTGCCACACTCGAAGACTTCTCTGTGCACCCCGAGTCCAGATTGTGA 461
Db 393 TCGCATCCAGCTGCAACCAATGTCGACCTTCACGTCATCCCGAGCCACCGTGGGTGA 452
QY 462 GGAGAAAGGGACAGCAGCTTTGAAATGCCACACCAAGGGCTTTCAGCCCCCATCATTTAC 521
Db 453 GGAGGGTGTGTGGCCCGCTTCAGTGCACAAATCCATGGGCTTCCCAACCCCTGATCAC 512
QY 522 TTGGGAAAGGACCAAGT---GACCGTGTGAGGAGCCCGGCTCATCATCTTCCCAA 578
Db 513 TTGGGAGAGACAGAGTCCCAATTGACACGGCAATGAGAGGTACACATTTGTCGCCAA 572
QY 579 GTGGCTCTCCAGATCTAGATGTCAGGACAGTGTATGAGGCTCTACCGCTCGGTGGC 638
Db 573 GGGGCTCTGCAGATCAGAGCTTCAGCTGAGGACGCTGGCATCTTCACTGTGTGGC 632
QY 639 CACCAATTGAGCCCGCAACGATTCAGCCAGGAGGCTCTGCTCACTGTGTGGCTTCAGAG 698
Db 633 CTCAAACATCGCCAGTATCCGGATCAGCCACGGGGCCAGGCTCACTGTGTGAGGCTCGGG 692
QY 699 GTCTTTGGAGGCTACGAGGGGACGATGTGTGTCATGTGTGAGGAGCCCGCCAGAGAACACAC 758
Db 693 CTCCTGGGGCTTACA-----AGGAGCCAGCCATCTCTGTGGGGCTTGAGAACCTCAC 743
QY 759 GGTAGTGTCTGGACAGAAATGATGATGAGTGTGGTGGCTCTGTGACCCCAACCCCTTT 818
Db 744 CTGACAGTGCACACCGCGTCTTGTAGTGTGTGCGCACGGGACCCCGGCCCAT 803
QY 819 TGTGCTCTGGGTCCTGACAGAGTGAAGGCTTATCTCCAGCGATGTGATC---GTTCTGGG 875
Db 804 TGTGCTCTGGAGCGCTGATGTTGGTGTGCTTATCGGGGTGGAGGCGATCCAGTGTGGG 863
QY 876 CCGAGCAATCTACTCATGCTCCAGCGCGGAGCCCTCGGCACTCTGTGAGTCTATGTCTGCG 935
Db 864 CACAGGAACCTCATCTATCTGACAGCTGAGCGTCCAGCACCTCTGGGCTTACGCTGTGC 923
QY 936 AGCCAAAGCCCTCAGCGGTGACITTCGCCACTGCGGCTGTGAGCTCGAGTGTCTGC 995
Db 924 AGCCAAAGACCTGGCACCGGGTGAAGAGACGGCACAGGGCCGCTGTGTGTGACG 983
QY 996 TGCCCCAGGACATCTCGCAGGACCCCGAGCGCTCTCGCGAGCGCGGCGCAGCACCGCGG 1055
Db 984 CCAGCTGAGTTGTGACGATCCCGAGTCCATCTCCAGGCCAGCTGGACACAGCCAT 1043
QY 1056 CTTGTGTGCGGGCTGCGGGAGCCACGCGCCGCGCTGCACTGTGCTGCAAGAGGAT 1115
Db 1044 GTTCACTGCAAGCCAGGGTGAGCCAGCGCTCATGTGCACTGTGCTGAAATGAGCA 1103
QY 1116 CCGTGTGACCAATGGGGCTCAAGTGCAGGGCGGTGGGGCGGCTGTGCTATCAC 1175
Db 1104 GTTGTGCGGCGAGGAGCCAGCTCAGGCTCAAGAAATAACAAGACACACTGACCAATTC 1163
QY 1176 TCAGATCGGCTCGCAGAGCGCTGGCTACTACAGTGCCTAGCAGAAACAGCGCGGGAAC 1235
Db 1164 TGAATCGGCTCGAGAGTGAAGCCATTTATAGTGTGTGGCGAGAACAGTGGGGCTC 1223
QY 1236 TGCTGTGCGGCTGCGCCCTGCGGTAGTGTGCGCGAGGGCTGCCAGCGGCCCGAC 1295
Db 1224 ATCACAGGCGAGTGCACCGGTGACCGTACTGTGGGCTGAGGGGCTCCCGGGCTCCCG 1283
QY 1296 TCGGFTCACAGCCAGCGCTGAGCAGCTCTCTGTGCTGTGTGGCTGTGGAGGGGCTGA 1355
Db 1284 CAATGTGCGGGCAGTCTCTGTGTCTTCCACTGAGGTGCGGTGTCTGTGAGTGAGCGGT 1343
QY 1356 GTTGCACAGCAGCAATCATTTGGCTTCTCTTCACTACCAAAAGGAGGGAGTGA 1415
Db 1344 GGCCAAACAGAGGATCATCGGCTACGTCCTGCAATCAGGAGGCTGTGACCCACC 1403
QY 1416 CAATGTGAGTACCACTTTGCAAGTAAACAAATGACACCAAGAGCTGCAGGTTGGGACCT 1475

Db 1404 GGAGCTGGAGTATCAGGAGGAGTACGCAAGAGCACCTTTGAGCACTGTGTGAGCACT 1463
QY 1476 GGAACCCCAACACGATATGATTTCTAGCTGTGGCTTACTCCCAAGCTTGGGGCCAGCG 1535
Db 1464 GGAGCCCTCCACAGCGCTACAGTTTCTACATCAAGGCTTACACACCAAGGGGCCAGCTC 1523
QY 1536 AACCTTCAGCCAGCCCTGTGTGCATACACTGACATGTCCCAGCGAGCACCCAG 1593
Db 1524 AGCCTCTGTGGCCACCTAGCTAGCACCTGTGGTGAAGCCCTGTGCCACCCCACTG 1581

RESULT 14

ABX71435

ID ABX71435 standard; cDNA; 3453 BP.

XX AC ABX71435;

XX DT 06-MAR-2003 (first entry)

XX DE Human cDNA encoding a novel tyrosine phosphatase, NHP4.

XX KW Human; ss; gene; chromosome 15; chromosome 3; tyrosine phosphatase;

XX KW Ig super family protein; gene therapy; NHP; novel human protein.

XX OS Homo sapiens.

XX PN US6465632-B1.

XX PD 15-OCT-2002.

XX PF 08-JUN-2001; 2001US-0877730.

XX PR 09-JUN-2000; 2000US-210607P.

XX PA (LEXI-) LEXICON GENETICS INC.

XX PI Walke DW, Scoville J, Turner CA, Friedrich G, Abuin A;

XX PI Zambrowicz B, Sands AT;

XX WPI; 2003-147071/14.

XX P-PSDB; ABU54203.

Novel isolated nucleic acid which encodes a novel human protein that shares sequence similarity with animal phosphatases, that is useful for generating antibodies, and as reagents in diagnostic assays

XX Claim 1; Column 31-36; 58pp; English.

The invention relates to an isolated nucleic acid molecule comprising a nucleotide sequence encoding a novel human protein (NHP) that shares sequence similarity with animal phosphatases (in particular tyrosine phosphatases and is a member of the Ig superfamily). The disclosed NHP polynucleotide sequences (from genes located on either chromosome 15 or chromosome 3) are useful for identifying coding sequences and in the identification of biologically relevant splice junctions. The NHP polynucleotide sequences are useful in gene therapy, for the diagnosis of mutant NHPs or inappropriately expressed NHPs, for the diagnosis of disease, for screening drugs effective in treatment of symptomatic or phenotypic manifestations of perturbing the normal function of NHP. Complementary sequences of the NHP polynucleotides used in conjunction with PCR to screen libraries, isolate clones and prepare cloning and sequencing templates. Labeled NHP nucleotide probes can be used to screen a human genomic library which is helpful for identifying polymorphisms, determining the genomic structure of a given locus/allele and designing diagnostic tests. The probe sequences also have use in defining and monitoring both drug action and toxicity. Oligonucleotides complementary to NHPs may encode or act as NHP antisense molecules, or may be used as part of ribozyme and/or triple helix sequences. The present sequence encodes an NHP.

XX Sequence 3453 BP; 979 A; 821 C; 769 G; 883 T; 1 other;

Query Match 5.1%; Score 315.4; DB 25; Length 3453;

Best Local Similarity 47.4%; Pred. No. 6.9e-71; Matches 1210; Conservative 0; Mismatches 1271; Indels 72; Gaps 6;			
QY	300	AGACAGCATGATGAGGAGCTCTTAGGATCTGGAGGTCACTGAGGCGAGCTATTCCTG	359
DB	279	ATACATGATGAGTGGAGCGGCGGAGGAGCGATCCGATGAAGATTTATTCAGTG	338
QY	360	TCTGGCCACAGCCCGCTAGGAGTGGTGGCCAGCGGTGTGTGTCAAGCTTGGCAC	419
DB	339	CTTGGCAATGAACAAATATGAGGCCATTTCTTAGTCAAAAAGCTCATCTTGCCTTATCAAC	398
QY	420	ACTCGAAGACTTCTCTGACCCCGAGTCCAGATTTGTGGAGGAACGGACAGCAGC	479
DB	399	TATTTCTGCAATTTGAAGTCCAGCCAAATTTCCACTGAGGTCCACGAAGGTGGATTTGCTCG	458
QY	480	CTTTGAATGCCACACCAAGGGCTTCCAGCCGCCCATCAATTAATCTGGGAAAGGACCAAGT	539
DB	459	ATTTCGATGCAAGATTTATCCACCCTCTCGAGTCAATACATGAGGATTCATTCGGAC	518
QY	540	GACCGTGGCTGAG---GAGCCCGCGCTCATCACTCTTCCCAAGTGGCTCTCCAGATCCT	596
DB	519	AACTCTACTATGACTATGAGCAGGATAACTGCGCTACCAACAGGAGTATTGCGATCTA	578
QY	597	AGATGTCAGGACATGATGAGGCTCCTACCGCTGCGTGGCCACCAATTCAGCCGCCA	656
DB	579	TGATGTCAGCCAAAGGGAATCTGGAAATTTATCGTTGTATTGCTGCCACTGTAGCCACCG	638
QY	657	ACGATTCAGCAGGAGGCTCGCTCACTGTGCGCTCAGAGGGTCTTTGGAGGTACACAG	716
DB	639	ACGTAAAGTATGGAGGCTCGCTAACTGTGATTCAG-----CTAAGGAGTCAAAATC	692
QY	717	GGGGCAGGATGTGTCTATTGGGAGCCCGGAGAAACACCGTAGTGTCTGGACAGAA	776
DB	693	CTTCCACACACCAACAAATTATAGCAGGTCACAGAAACATACAACTCTTTCATCAGAC	752
QY	777	TGTAGTATGAGTGGTGGCTCTGTGACCCCAACCCCTTTTGTCTCGGTCCGACA	836
DB	753	TGTAGTTTGGAAATGCATGGGCCACAGGAAATCCCAACCAATCAATTTCTTGAGCGGCT	812
QY	837	GGATGGAAGCTAT---CTCCAGGATGTCACTCGTTCTGGCCGGACCAATCTACTCAT	893
DB	813	TGATCACAAATCCATTTGATGTCTTAACTCGGGTACTTGGAAATGTGAATCTCATGAT	872
QY	894	CGCCAGCGCAGCGCTCGGCACTCTGGAGTCTATGTCGCCAGCAACCAAGCCCTCAC	953
DB	873	ATCTGATGTCAGGCTAGACATGCTGGAGTATATGTTTGTGGGCCACTACCCCTGGCAC	932
QY	954	CGTGACTTCGCACTGGCGCTGTGAGCTCGAGTGTGCTGCCCGCCAGCCATCTGCA	1013
DB	933	ACGCAACTTTACAGTTGTATGGCAACTTTAACTGTATTAGCTCTCTTCAATTTGTGA	992
QY	1014	GGCACCAGGCGCTCTGGCGACCGGGCCAGCACCGCGGCTTCGTGTGCGGGCGTC	1073
DB	993	ATGGCCAGAAAGTTTAAACAGGCTCTGAGCTGGCACTGCTCGAATTTGTGTGAGGCGACA	1052
QY	1074	CGGGAGGACACGGCCCGCTGCACTGTGCTGCACACCGGATCCCGTTGGACCCAAATGG	1133
DB	1053	AGGAATCCCTCTCCAAAGATGTATGGTTGAAATGAAGGAAGATACATTCGAAATGG	1112
QY	1134	CGCGCTCAAGGTGACAGGCGGTGGCGAGCTTGTGTCATCACTCAGATCGGCTGCGAGGA	1193
DB	1113	TAGAATTAATATGTACACAGT-----AAATTGGTAATTAACACAGATTTATTCGTGAAGA	1166
QY	1194	CGCTGGCTACTACAGTGGTAGCAAGAAACAGCGCGGGAACTGCTGTGCGGTGCGCC	1253
DB	1167	TGATGCTATTTATCAGTGTGATGGCTGAGAAATAGCCAAAGGATCTATTTTATCTAGAGCCAG	1226
QY	1254	CTGCGGTGATGTTGGCGAGGGGCTCCAGCGCCCGGACTCGGGTCAAGCCACGCC	1313
DB	1227	ACTGACTGTAGTGTGATGACAGAGACAGACCCAGTCTCCCTATATATGTACATGCTGAAAC	1286
QY	1314	GCTGAGCAGCTCTCTGTGCTGGTGGCTGGGAGCGGCTGTAGTTGCAAGGAGCAAAAT	1373

DB	1287	CATGTCAGGCTCAGCCATTTCTTTAGCCTGGGAGAGGCCACTTTATTAATTCAGACAAAGT	1346
QY	1374	CAITGGCTTCTCTCTTCACTACCAAAAGGCAAGGGAGTGACAAATGTGGAGTACCAGTT	1433
DB	1347	CAITGGCTTCTCTCTGTA CACTACATGAAGCAGAGGTTTAAATTAATGAAGATTCAGT	1406
QY	1434	TGCAGTAAACAAATGACACACAGAGCTGCGAGTTGCGGACCTGGAACCCAAACAGGATTA	1493
DB	1407	AGTCATCGGAAATGACACAACTCATTTATTTATTTGATGACTTAGAGCCTGCGCAGCAATTA	1466
QY	1494	TGAGTTCTAGTGGTGGCTACTCCAGCTGGGGGGCAGCGAACCTCCAGGCCAGCCT	1553
DB	1467	TACTTTCTTACATTTGTAGCATATATGCCAATGGGAGCAGCAGATGTCGACCATGTGAC	1526
QY	1554	GGTCATACACTGACGATGTCCCGCAGCGCAGCACCCAGCTTACCTTGTTCAGCCCCCAA	1613
DB	1527	ACAGAACTACTAGAGGATGTTCCTCGAGACCTCTGGAATTAATTTGACAACTGCGAAG	1586
QY	1614	CCCCCTGGACATCAGGCTGGCATGCTGCCCCCTGCTCCAGCTCGAGCAATGACAGGT	1673
DB	1587	TCCCACTGATATCTCATCTCTGGCTGCCAATCCAGCCAAATATCGCGGGGCCAAGT	1646
QY	1674	GCTGAAGTACAAATAGATAGGTACGGTTGGGGAAGGAGATCAGGTTTCTTCTCAAGGCT	1733
DB	1647	GGTCTGTATGCTGTCTTTCGCGCTAAAGTACTGAGAAITTCAAATCCAAAGTTCTGGAGCT	1706
QY	1734	GCCTGGAAATGAGACACAACTTACGTTAAACTCACTTCAGCCAAACAAAGTTGACCGAGT	1793
DB	1707	CCCGGGAGCAGCATGATGATGATCTTTTGGAAAGGCTGAACTGACAGTGTCTACCTGCT	1766
QY	1794	CCGGAATTCAGCTGGCACTGGCGCTATGGAGTCCCTTCTCAGTGGATGACGACACAG	1853
DB	1767	TCGGATTACTGCTGCCACAGAGTGGGGCTGGGAGAGTCACTAGTATGGACTTCACATAG	1826
QY	1854	GACACTGCTGTGCACAAACACAGAGCATGTTCCCTTTGCCCCCTCGCAGAAITTTGAAGGTGAG	1913
DB	1827	GACGCCCAAAGCTACAAGCGTGAAGCCCTTAAGTCTCAGAGTTGCAATTTGGAGCTCT	1886
QY	1914	GGCAAGATGAGTCCCTGCTGTGTGATGCGAGCGGCCCTCACCCACCCAGCATCTC	1973
DB	1887	GAACTGTACCACTATTTCTGTGAGTGGCAGCAAGATGTAGAGGACACAGCTGTATTCA	1946
QY	1974	TGGATACAAACTCTACTGGGAGAGGTGGGAAACAGAGAGGAGGAGGAGTGTGACCGGCC	2033
DB	1947	GGGCTACAAGCTGTACTACAAGGAAGAGGGCAGCAGGAGAA-----1988	
QY	2034	CCCAGGGGCTCGTGGAGATCAAGCTTGGGAAGTGGGCGCGGCCCTGAGAGAGAAAGT	2093
DB	1989	-----TGGGCCCATTTTCTTGGATACCAAGGA 2015	
QY	2094	GAAGCAGTATGAACCTGACCCAGTTAGTCCCTGGCAGGCGGTACGAGGTGAAGCTCGTAGC	2153
DB	2016	CCTACTCTATCTCTCAGTGGCTTAGACCCGAGAGAAATATCATGTGAGACTCCTGCG	2075
QY	2154	TTTCAACAAACACAGGACCGCTACGCTGTGTGTGGAAAGGCGCAAGACGAGAGGCGCC	2213
DB	2076	TTACAAACACATAGACGATGGCTATCAGGACAGATCAGACTGTCTCAGCAC---TCCAGGATG	2132
QY	2214	CACGCGACAGCTGCTATTCAGAGGGGGCCACCGCTGCTCTTGCCTCATGTGCCCATGTCCAGCAGA	2273
DB	2133	CGTGTCTGTTGCTGATGCGCATGCTCCCTCTCCACCAACCCCAACCATCTCTATGCGAA	2192
QY	2274	GTCAACAGCTCCACTTCCATTTGGCTTGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2333
DB	2193	GGCTAAACACCTCATCTTCCATCTTCTGCACTGAGAGGAGGCTGCAATTCACCGCTGCACA	2252
QY	2334	GATTTGTAACCTACACTGTAGCGCTTCCGCCCTTGGGGCTCAGGAATGCTTCCCTGCTCAC	2393
DB	2253	AATCATTTAACTACACCATCCGCTGTATCTCTGTTGGCTGAGAAATGCTTCTTGGTCT	2312
QY	2394	CTACTATACAGCTCTGAGAGAGCATTTCTCATTTGGCGGCTGAAACCAATTTACCAAGTA	2453
DB	2313	GTACCTTCAACATCAGAAACTCACAATGTTGTTCAAGTCTTAGAACCCAAACACCAATA	2372

Db 922 AGGCTCAAGAATAACACAGCACACTGACCATTTCTGGAATCGGTCTCTGAGGATGAAGCC 981
Qy 1201 TACTACCAAGTGGTAGCAGAAACACAGCGCGGGAJACTGCCTGTGTGCGCTGCGCCCTTGGCG 1260
Db 982 ATTTATCAGTGTGTGGCGGAGAACAGTGGGGCTCATCACAGGCCAGTGGCCAGGCTGACC 1041
Qy 1261 GTAGTGGTGGCGGAGGGGCTGCCAGCGCCCGGACTCGGCTCACAGCCACGCCCGCTGAGC 1320
Db 1042 GTACTGTGGGTGAGGGGCTCCCCGGGCTCCCCCGCAATGTGCGGGCAGTCTCTGTGTCT 1101
Qy 1321 AGCTCCTCTGTGCTGGTGGCTTGGGAGCGGCTGTGAGTTGCACAGCGAGCAAAATCAATTGGC 1380
Db 1102 TCCACTGAGGTGGGTGTGTCTGTGAGTGAGCGGTGGCCAAACACCAAGGAGTCAATCGGC 1161
Qy 1381 TTCTCTCTTCACTACCAAGGCAAGGGGAGTGGACAAATGTGGAGTACCAAGTTTGCAGTA 1440
Db 1162 TACGTCTCTGCACATCAGGAAGGCTGTGACCCACCGGAGCTGGAGTATCAGGAGGCAGTC 1221
Qy 1441 AACATGACACACACAGAGCTGAGGTTGGGACCTGGGACCTGGAAACCAACACGGATTATGAGTTC 1500
Db 1222 AGCAAGAGCACCTTTTCAGCACCTGGTTCAGCGACCTGGGAGCCCTCCACAGCCTACAGTTTC 1281
Qy 1501 TACGTGGTGGCTACTCCAGCTGGGGGCCAGCCGAACTCCAGCCAGCCCTGTGTGCAT 1560
Db 1282 TACATCAAGGCTACACACCAAGGGGGGCCAGCTCAGCCTCTGTGCCCCACCCCTAGCTAGC 1341
Qy 1561 ACCTGGACGATGTCCCGAGCGCAGCACCCCGAG 1593
Db 1342 ACCCTGGGTGAAGCCCTGCCCCACCCCACTG 1374

Search completed: October 14, 2003, 10:17:18
Job time : 1181.58 secs

121 DB CTGCAAGTGATCTCTGGGCTGAGCAGGCTGTGTGCTGCACTGCACTTTTGGGGCTACA 180
181 QY GCTGCTGGGCTCCGACCAAGGTGACATGGAGCAAGGATGGAGACTGTACTAGAGCAT 240
181 DB GCTGCTGGGCTCCGACCAAGGTGACATGGAGCAAGGATGGAGACTGTACTAGAGCAT 240
241 QY GAGAACTGACCTGCTACCAATGCTCCCTGTGTGCTGTCTCACCCTTAGAGCAAGAA 300
241 DB GAGAACTGACCTGCTACCAATGCTCCCTGTGTGCTGTCTCACCCTTAGAGCAAGAA 300
301 QY GACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTCTGT 360
301 DB GACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTCTGT 360
361 QY CTGGCCCAAGCCCGCTAGAGTGTGGCCAGCCAGGTTGCTGTGTGCTGAAGCTTGGCCA 420
361 DB CTGGCCCAAGCCCGCTAGAGTGTGGCCAGCCAGGTTGCTGTGTGCTGAAGCTTGGCCA 420
421 QY CTGGAAGACTTCTCTGCAACCCGAGTCCAGATTGTGAGGAGAAACGGGACAGCACGC 480
421 DB CTGGAAGACTTCTCTGCAACCCGAGTCCAGATTGTGAGGAGAAACGGGACAGCACGC 480
481 QY TTTGAAATGCCACCAAGGSCCTTCCAGGCCCTCATATTACTTTGGGAAAGGACCAAGGTG 540
481 DB TTTGAAATGCCACCAAGGSCCTTCCAGGCCCTCATATTACTTTGGGAAAGGACCAAGGTG 540
541 QY ACCGTGCTGAGAGCCCGGCTCATCTCTTCCCAAGTGGCTCTCCAGATCTTAGAT 600
541 DB ACCGTGCTGAGAGCCCGGCTCATCTCTTCCCAAGTGGCTCTCCAGATCTTAGAT 600
601 QY GTCCAGGACATGATCAGGCTCTACCGCTGCTGCGCCACCAATTCAGCCCGCCAAACGA 660
601 DB GTCCAGGACATGATCAGGCTCTACCGCTGCTGCGCCACCAATTCAGCCCGCCAAACGA 660
661 QY TTCAGCCAGAGGCTCGCTCACTGTGGCCCTCAGAGGCTCTTTGGAGGCTACAGGGGG 720
661 DB TTCAGCCAGAGGCTCGCTCACTGTGGCCCTCAGAGGCTCTTTGGAGGCTACAGGGGG 720
721 QY CAGGATGTGTGCTATTGTGGAGCCCGCAGAGAACACACCGTAGTGTCTGGACAGATGTA 780
721 DB CAGGATGTGTGCTATTGTGGAGCCCGCAGAGAACACACCGTAGTGTCTGGACAGATGTA 780
781 QY GTGATGAGTGTGCTGCTGCTGACCCACCCCTTTGTGTCTGCTGCTGCTGCTGCTGCT 840
781 DB GTGATGAGTGTGCTGCTGCTGACCCACCCCTTTGTGTCTGCTGCTGCTGCTGCTGCT 840
841 QY GGAAAGCCTATCTCCAGGATGTATCGTTCTGGGCGGACCAATCTACTCATCGCCAGC 900
841 DB GGAAAGCCTATCTCCAGGATGTATCGTTCTGGGCGGACCAATCTACTCATCGCCAGC 900
901 QY GCGAGGCTCGGCACTCTGAGTGTATGTCTGCGGAGCCAAAGGCCCTCTACGCGTGAC 960
901 DB GCGAGGCTCGGCACTCTGAGTGTATGTCTGCGGAGCCAAAGGCCCTCTACGCGTGAC 960
961 QY TTCGCCACTCGGGCTGTGAGCTCCGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
961 DB TTCGCCACTCGGGCTGTGAGCTCCGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
1021 QY GAGCGCTCTCGGGAGCGGGGCGCAGCACCGCGCTTCTGTGTGCTGCTGCTGCTGCTGCTGCT 1080
1021 DB GAGCGCTCTCGGGAGCGGGGCGCAGCACCGCGCTTCTGTGTGCTGCTGCTGCTGCTGCTGCT 1080
1081 QY CCAAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
1081 DB CCAAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
1141 QY AAGGTGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
1141 DB AAGGTGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
1201 QY TACTACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
1201 DB TACTACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260

1201 DB TACTACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
1261 QY GTAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
1261 DB GTAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
1321 QY AGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
1321 DB AGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
1381 QY TTCTCTCTTCTACTACTACCAAGGCAAGGGAGTGGCAATGTGGAGTACAGGTTTGCAGTA 1440
1381 DB TTCTCTCTTCTACTACTACCAAGGCAAGGGAGTGGCAATGTGGAGTACAGGTTTGCAGTA 1440
1441 QY AACAAATGACACACAGAGCTGAGGTTCGGGACCTGGAACCCCAACCGGATTTATGAGTTC 1500
1441 DB AACAAATGACACACAGAGCTGAGGTTCGGGACCTGGAACCCCAACCGGATTTATGAGTTC 1500
1501 QY TACGTGTGTGGCTTACTCTCCAGCTGCGGCGCAGCGCACTCCAGCCAGCCCTGCTGAT 1560
1501 DB TACGTGTGTGGCTTACTCTCCAGCTGCGGCGCAGCGCACTCCAGCCAGCCCTGCTGAT 1560
1561 QY ACACCTGGAAGATGCTCCCGAGCGCAGCACCCAGCTTACCTTGTCCAGCCCAACCCCTCG 1620
1561 DB ACACCTGGAAGATGCTCCCGAGCGCAGCACCCAGCTTACCTTGTCCAGCCCAACCCCTCG 1620
1621 QY GACATCAGGTGGCATGCTGCCCTGCGCTCCAGCTTACCTTGTCCAGCCCAACCCCTCG 1680
1621 DB GACATCAGGTGGCATGCTGCCCTGCGCTCCAGCTTACCTTGTCCAGCCCAACCCCTCG 1680
1681 QY TACAAGATAGAGTACGGTTTGGGGAAGGAAGATCAGGTTTCTCCACAGGTGCTGGA 1740
1681 DB TACAAGATAGAGTACGGTTTGGGGAAGGAAGATCAGGTTTCTCCACAGGTGCTGGA 1740
1741 QY AATGAGACAACTTACGTTTAACTCACTTCAGCCCAACCAAGTGTACCGAGTCCGAT 1800
1741 DB AATGAGACAACTTACGTTTAACTCACTTCAGCCCAACCAAGTGTACCGAGTCCGAT 1800
1801 QY TCAGCTGCGCAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
1801 DB TCAGCTGCGCAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
1861 QY GGTGTGCAACACAGAGCCATGTTCCCTTTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 1920
1861 DB GGTGTGCAACACAGAGCCATGTTCCCTTTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 1920
1921 QY ATGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
1921 DB ATGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
1981 QY AAACTCTACTGCGGAGAGGTGGGAACAGAGGAGGAGGAGATGCTGACCGCCCGCCAGGG 2040
1981 DB AAACTCTACTGCGGAGAGGTGGGAACAGAGGAGGAGGAGATGCTGACCGCCCGCCAGGG 2040
2041 QY GGTGTGCAACACAGAGCCATGTTCCCTTTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 2100
2041 DB GGTGTGCAACACAGAGCCATGTTCCCTTTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 2100
2101 QY TATGAACTGACCCAGTGTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
2101 DB TATGAACTGACCCAGTGTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
2161 QY AAAACAGAGAGCGGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
2161 DB AAAACAGAGAGCGGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
2221 QY GACCTGCTTATCCAGAGGGGGCCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
2221 DB GACCTGCTTATCCAGAGGGGGCCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
2281 QY AGCTTCACTTCCATTTGGCTTGGTGGAGAGAGCCAGATTTTACCACTGCTGCTGCTGCTGCTGCT 2340
2281 DB AGCTTCACTTCCATTTGGCTTGGTGGAGAGAGCCAGATTTTACCACTGCTGCTGCTGCTGCTGCT 2340


```
Db 4501 |||||CTTGTCTTATGTCGTGTTCTCGTCCCTGAGTTCAACCTATGCACTTCTTCTTAACA 4560
Qy 4561 ACATGACTACCTCATGTCGTCTTACACATAGTGTGACCCCTGGGTCCCAACAGTCC 4620
Db 4561 ACATGACTACCTCATGTCGTCTTACACATAGTGTGACCCCTGGGTCCCAACAGTCC 4620
Qy 4621 CTGCCAACCGCTTCTCGGCGCAGATGAGCCCACTCCCAAGTAGATCTGGAAGAGACCCCTG 4680
Db 4621 CTGCCAACCGCTTCTCGGCGCAGATGAGCCCACTCCCAAGTAGATCTGGAAGAGACCCCTG 4680
Qy 4681 TGGCTTGTCTGGCTGCCCTCCCTCGTGTGTGAGATGAGAAGTCTTCTATGGAAGAGAT 4740
Db 4681 TGGCTTGTCTGGCTGCCCTCCCTCGTGTGTGAGATGAGAAGTCTTCTATGGAAGAGAT 4740
Qy 4741 GAGTCCAGGCTGCACAGGGGMAACCCCAAGAGGGGTAGGGAGTGAACCAAGAGGCTGA 4800
Db 4741 GAGTCCAGGCTGCACAGGGGMAACCCCAAGAGGGGTAGGGAGTGAACCAAGAGGCTGA 4800
Qy 4801 AAAAAAATGGCTGCCACCCATCTGCACAGAGAGATGGGTGTGTCTTTTGAAGTGCAGTC 4860
Db 4801 AAAAAAATGGCTGCCACCCATCTGCACAGAGATGGGTGTGTCTTTTGAAGTGCAGTC 4860
Qy 4861 CTGGCTGAACCTGAAGGGGTGAGGAGGGGAGCTACTGGGGCTGCCATGGCTCAGTCC 4920
Db 4861 CTGGCTGAACCTGAAGGGGTGAGGAGGGGAGCTACTGGGGCTGCCATGGCTCAGTCC 4920
Qy 4921 CTGACCCCTGGAGCCCTGAACCTGGCTTACAGAGATGGGTGTGTCTTCTCAAGATGCTGT 4980
Db 4921 CTGACCCCTGGAGCCCTGAACCTGGCTTACAGAGATGGGTGTGTCTTCTCAAGATGCTGT 4980
Qy 4981 AAGGGAAGTCTTTGATAGGAAGAGGGCGCTGGCTCATTTTATTTTATCTTCTTTTACA 5040
Db 4981 AAGGGAAGTCTTTGATAGGAAGAGGGCGCTGGCTCATTTTATTTTATCTTCTTTTACA 5040
Qy 5041 CTGAATCCCAAAATCATCTTTACCAAAAGGGCCCAAGCTGAGTGTATTTCTGAGTCAC 5100
Db 5041 CTGAATCCCAAAATCATCTTTACCAAAAGGGCCCAAGCTGAGTGTATTTCTGAGTCAC 5100
Qy 5101 AAGAGCCATGCCATCTCTCTGGTTCTCAGCTCAGTCATGTCAGAGATGTCAGATCCA 5160
Db 5101 AAGAGCCATGCCATCTCTCTGGTTCTCAGCTCAGTCATGTCAGAGATGTCAGATCCA 5160
Qy 5161 GTGGCATCTGCTCTTCTGTCACATCTTTCTATTTTCAACTGGCTGGCACAATCAAGTGT 5220
Db 5161 GTGGCATCTGCTCTTCTGTCACATCTTTCTATTTTCAACTGGCTGGCACAATCAAGTGT 5220
Qy 5221 AACTCTGGCTTCTGGGCCAAAGTTAGAAATAACCAAGTCTATTTTCCCTTTATTTTATTA 5280
Db 5221 AACTCTGGCTTCTGGGCCAAAGTTAGAAATAACCAAGTCTATTTTCCCTTTATTTTATTA 5280
Qy 5281 TTTTATTTTATTTATGTCCTTTCAGTGGAGTTGTAGCTTCTGAAAGGCTGTGTTTATTT 5340
Db 5281 TTTTATTTTATTTATGTCCTTTCAGTGGAGTTGTAGCTTCTGAAAGGCTGTGTTTATTT 5340
Qy 5341 AGCTCTGTGTGCTCACTCATGTTTGAACCCCAACCACTTCTCTCTCCCTCTTCAGC 5400
Db 5341 AGCTCTGTGTGCTCACTCATGTTTGAACCCCAACCACTTCTCTCTCCCTCTTCAGC 5400
Qy 5401 CAGCTATGATAACACATTAAGATTTATTAATGCTGGCTTCGTATCTCAATTAAGACAGGAT 5460
Db 5401 CAGCTATGATAACACATTAAGATTTATTAATGCTGGCTTCGTATCTCAATTAAGACAGGAT 5460
Qy 5461 TGTCACTTGAACCTACTTCTATAGCATTTCAAAGTGGCCACCGGCCAAACCAACCGTATGTTT 5520
Db 5461 TGTCACTTGAACCTACTTCTATAGCATTTCAAAGTGGCCACCGGCCAAACCAACCGTATGTTT 5520
Qy 5521 CTTTCATGCTCTGAAGTCTCAAGAGCTCATTTTGTGTTTCTGTTAGATCTTTTCTCTCC 5580
Db 5521 CTTTCATGCTCTGAAGTCTCAAGAGCTCATTTTGTGTTTCTGTTAGATCTTTTCTCTCC 5580
Qy 5581 TTGCTTGAATGAATAACCGTTTAAACAGTAGGCTCTTAGCATCACACACATAGTCAT 5640
|||||
```

```
Db 5581 TTGCTTGAATGAATAACCGTTTAAACAGTAGGCTCTTAGCATCACACCATAGTCAT 5640
Qy 5641 TCCTCATGTTCTTGTGTTTAAACAGCACTTGAAGTCTCGGTTTAAATTAATAGCTGCAAA 5700
Db 5641 TCCTCATGTTCTTGTGTTTAAACAGCACTTGAAGTCTCGGTTTAAATTAATAGCTGCAAA 5700
Qy 5701 TGAGCAATTTATATACCCATTTAGGCTGGGTGGAAATTTGTTCTCAAAGCAAAATAGTAA 5760
Db 5701 TGAGCAATTTATATACCCATTTAGGCTGGGTGGAAATTTGTTCTCAAAGCAAAATAGTAA 5760
Qy 5761 TAAATCTGGTATCTGCCCTATAACTCACAGTTGTATGAAGAGTAGCCAGACTCATACGA 5820
Db 5761 TAAATCTGGTATCTGCCCTATAACTCACAGTTGTATGAAGAGTAGCCAGACTCATACGA 5820
Qy 5821 TTATATATGATTTGGGGTTCTGAGTAACTGGGAGGTGTAGCTTTGTGACTTTGTAGCACC 5880
Db 5821 TTATATATGATTTGGGGTTCTGAGTAACTGGGAGGTGTAGCTTTGTGACTTTGTAGCACC 5880
Qy 5881 AGGCTCTTATAGGAAGTCTGTGTGGCTTTTACAGGGCAATAGTCCCTTGTGCTTTGCC 5940
Db 5881 AGGCTCTTATAGGAAGTCTGTGTGGCTTTTACAGGGCAATAGTCCCTTGTGCTTTGCC 5940
Qy 5941 ATGGATGCTTAAAGTCTTGTGGAGTCTCATTTAAGAAATTCCTTTCTCGAAGCATGACAA 6000
Db 5941 ATGGATGCTTAAAGTCTTGTGGAGTCTCATTTAAGAAATTCCTTTCTCGAAGCATGACAA 6000
Qy 6001 GTGTATCGCAATCTTACATGCTCACTGTTTACCTGCTAGTTTGTGCTGGGTATTTT 6060
Db 6001 GTGTATCGCAATCTTACATGCTCACTGTTTACCTGCTAGTTTGTGCTGGGTATTTT 6060
Qy 6061 AATTGCACCTTCCAGCATCTGCTTCTTACAAATATGATATTTTATTTGTATACAC 6120
Db 6061 AATTGCACCTTCCAGCATCTGCTTCTTACAAATATGATATTTTATTTGTATACAC 6120
Qy 6121 TAAAGTGTGTGATCATGTATCTGCTGCTGTAAGAAATTAATAACTATTTTTCAGAC 6176
Db 6121 TAAAGTGTGTGATCATGTATCTGCTGCTGTAAGAAATTAATAACTATTTTTCAGAC 6176
```

RESULT 2

```
US-09-754-997A-45
; Sequence 45, Application US/09754997A
; Patent No. US20020102551A1
; GENERAL INFORMATION:
; APPLICANT: Salbaum, Michael J.
; TITLE OF INVENTION: NOPE Polypeptides, Encoding Nucleic
; FILE OF INVENTION: Acids and Methods of Use
; FILE REFERENCE: P-NI 4552
; CURRENT APPLICATION NUMBER: US/09/754,997A
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: US 60/174,496
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: US 60/205,789
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 3756
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-754-997A-45
```

```
Query Match 60.8%; Score 3756; DB 10; Length 3756;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3756; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGCGGCGCGGACACGCGGCGCGGCTCTCGGTGCTGACCTTCTGCTGCTGCTCGCG 60
Db 1 ATGGCGCGGCGGACACGCGGCGCGGCTCTCGGTGCTGACCTTCTGCTGCTGCTCGCG 60
Qy 61 CGCGGGGAGCTGCTCATTTGCCCCCAGGAGACAACTGTCAAGCTGAGCTGTGATGAGGACCC 120
Db 61 CGCGGGGAGCTGCTCATTTGCCCCCAGGAGACAACTGTCAAGCTGAGCTGTGATGAGGACCC 120
```


QY 183 TGCTGGGCTCGACACAGGCTGACATGAGCAGAGATGGAAGACATGTACTAGAGACATGA 242
Db CGCTGGACCCCCACACAGGCTGACCTGGAGCAAGATGGGAGACACCTGCTGGAGACGA 248
QY 243 GAACTGCACTGCTACCAATGGCTCCCTGCTGGCTGTCTCACCCTAGAGCAAGAGA 302
Db CCATTACACCTGCTGCGCAATGGTTCCTGCTGGCTGTCCAGCCACTAGACCCCAATGG 308
QY 303 CAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCCTGTCT 362
Db CAG---TGACGAGTCAGTCCCTGAGGCTGTGGGGTCAATTGAAGCACTATTCGTGCT 365
QY 363 GGCACACAGCCGCTAGAGATGGTGGCCAGCCAGAGTTGCTGTGTGCTGAAGCTTGCACACT 422
Db AGCCACAGCCGCTCGAGTGTGCGCCAGCCAGACTGTCTGTCTGAAGCTTGCACACT 425
QY 423 CGAGACTTCTCTGCAACCCGAGTCCAGATTGTGGAGAGAACGGGACAGCAGCTT 482
Db CGCAGACTTCTCTGCAACCCGAGTCTCAGACGCTGAGGAGAACGGGACAGCTCGCTT 485
QY 483 TGAATGCCACCAAGGGCTTCCAGCCCTCATCTTACTTTGGGAAAGGACCAAGGTGAC 542
Db TGAGTGCCACATTGAAGGCTGCCAGCTCCCATCTTACTTTGGGAAAGGACCAAGGTGAC 545
QY 543 CGTGCCTGAGGAGCCCGCTCATCACTTTCCCAAGTGGCTCTCTCGAGATCTTAGATGT 602
Db ATTGCTGAGGAGCTCGGCTCATGTGCTTCCCAACGGGCTCTCTCAGATCTCTGGATGT 605
QY 603 CCAGACAGATGATGAGGCTCTACGCTGGTGGCCAGCCAAATTCAGCCCGCCCAAGATT 662
Db TCAGGAGATGATGAGGCTCTACGCTGGTGGCCAGCCAACTCAGCTCGCCAGCACTT 665
QY 663 CAGCCAGAGGCTCGCTCACTGTGGCTCTCAGAGGCTCTTTGAGGCTTACAGGGGCA 722
Db CAGCCAGAGGCTCGCTCACTGTGGCTCTCAGAGGCTCTTTGAGGCTTACAGGGGCA 725
QY 723 GGATGTGTCTATTTGGCAGCCCGCCAGAGAACACCAAGTGTGTCTGGAAGATGTAGT 782
Db GGAAGTGTCTATTTGGCAGCCCGCCAGAGAACACCAAGTGTGTCTGGAAGATGTAGT 785
QY 783 GATGGAGTGTGGCTCTGTGACCCCAAGCTTTGTGTCTGGGTCCAGAGGATGG 842
Db GATGGAATGTGGCTCTGTGACCCCAAGCTTTGTGTCTGGGTCCAGAGGATGG 845
QY 843 AAAGCTATCTCCAGGATGTCTGCTGGCCGAGCAATCTTACTCATCGCCAGGCT 902
Db GAAGCCCATCTCCAGAGATGTCTGCTGGCCGAGCAATCTTACTCATCGCCAGGCT 905
QY 903 GCAGCTCGGCACTGTGAGTCTATGTCTGCGAGCCCAAGCCCTCTACGCTGACTT 962
Db GCAGCTCGGCACTCGGCTCTATGTCTGCGGCGCCAAAGCCCGCCAGCGGACTT 965
QY 963 CGCCACTGGGCTGTGAGTCTCGAGTGTCTGTGCTGGCCAGGCTCTGCGAGGACCCGA 1022
Db CGCCACTGAGCTGTGAGTCTCGGCTCTGCTGCTGGCCAGGCTCTGCGAGGACCCGA 1025
QY 1023 GGCCTCTCGGAGCGGGCCAGCACCGCGCTTCTGTGTGCGGGCTCGGGGAGCC 1082
Db GGCCTCTCGGAGCGGGCCAGCACCGCGCTTCTGTGTGCGGGCTCGGGGAGCC 1085
QY 1083 ACGGCCCGCTGCACTGGCTGCAAGCGGATCCGTTGCGAACCCCAATGGGCGCGTCAA 1142
Db GCGGCCAGCGCTGCGCTGCAAGCGGCGCGCTGCGGGCCCAAGCGGCGGCTCAA 1145
QY 1143 GGTGAGGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1202
Db GGTGAGGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1205
QY 1203 CTACAGTGTGCTGCAAGAACGCGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1262
Db CTACAGTGTGCTGCAAGAACGCGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1265

QY 1263 AGTGTGGCGAGGGGCTGCCACAGGCCCCGACCTCGGGTCAAGCCACGCGCTGAGCAG 1322
Db GGTGTGGCGAGGGGCTGCCACAGGCCCCGACCTCGGGTCAAGCCACGCGCTGAGCAG 1325
QY 1323 CTCCTCTGTGTGTGTGGAGCGGCTGTGAGTGTGACAGCGAGCAAAATCATTTGGCTT 1382
Db CTCCTCTGTGTGTGTGGAGCGGCTGTGAGTGTGACAGCGAGCAAAATCATTTGGCTT 1385
QY 1383 CTCCTTTCACTACCAAAAGGCAAGGGAGTGGACAATGTGTGAGTACCAAGTTTGCAGTAAA 1442
Db CTCCTCCACTTACCAGAAGGCAAGGGAGTGGACAATGTGTGGAATACCAAGTTTGCAGTAA 1445
QY 1443 CAATGACACACACAGAGCTGCAAGTTTGGGACCTGGAAACCCAAACGGAATTAATGATTTCTA 1502
Db CAACGACACACACAGAGCTTACAGGTTTGGGACCTGGAAACCCAAACAGATTAATGATTTCTA 1505
QY 1503 CGTGTGGCTTACTTCCAGCTTGGGGCCAGCGAACTTCCAGCCCGCCCTGTGTGCTATAC 1562
Db CGTGTGGCTTACTTCCAGCTTGGGGCCAGCGAACTTCCAGCCCGCCCTGTGTGCTATAC 1565
QY 1563 ACTGAGCAGATGTCCCCAGCGCAGCACCCCGCTTACTTGTCTCAGCCCGCCCAACCCCTCGGA 1622
Db ACTGAGTGTTCCTCCAGTGCAGCACCCCGCTTCTCTCTCAGCCCGCCCAACCCCTCGGA 1625
QY 1623 CATCAGGGTGGCATGGCTGCCCCCTGCTCAGCTCAGCAATGGAAGTA 1682
Db CATCAGGGTGGCTGCTGCCCCCTGCTCAGCTCAGCAATGGAAGTA 1685
QY 1683 CAGATAGATGATGCTTTGGGAGGAGATCAGGTTTCTTCCACCGAGGCTCTGGA 1742
Db CAGATAGATGATGCTTTGGGAGGAGATCAGGTTTCTTCTACTGAGGTCGAGGAA 1745
QY 1743 TGAGACAACTTACGTTAAACTCACTTTCAGCCAAACAAAGTGTACCGAGTCCGAGTTTC 1802
Db TGAGACAACTTACGTTAAACTCACTTTCAGCCAAACAAAGTGTATCGAGTACCGAGTTTC 1805
QY 1803 AGCTGGCACTGGGCTGCTGTATGAGTGTCTTCTCAGTGTGATGAGCAAGCACTGG 1862
Db GGCTGTGTACAGCAGCGGCTTGGGGCTTCCAGTGTGATGATGAGCAAGCACTGG 1865
QY 1863 TGTGCAAAACAGAGGCTGTTTCTTTCCTTGGCTGAGGATGAGGTCGAGGCAAGAT 1922
Db TATGCAAAACAGAGGCTGTTTCTTTCCTTGGCTGAGGATGAGGTCGAGGCAAGAT 1925
QY 1923 GGAGTCTCTGTGTGTGATGAGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1982
Db GGAGTCTCTGTGTGTGATGAGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1985
QY 1983 ACTCTAATGGGAGGAGTGGGAAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2042
Db ACTATAATGGGAGGAGTGGGAAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2045
QY 2043 TCGTGGAGATCAAGCTTGGGAGCTGCGGGCTGCGGCTGAGAGGAGGAGGAGGAGGAGGAGG 2102
Db CCGTGGAGATCAAGCTTGGGAGCTGCGGGCTGCGGCTGAGAGGAGGAGGAGGAGGAGGAGG 2105
QY 2103 TGAATCAACCACTTGT 2162
Db TGAATCAACCACTTGT 2165
QY 2163 ACAGAGAGCGGCTACGCTGT 2222
Db ACATGAGGATGCTATACGAGCAGT 2225
QY 2223 CCGTGTCTATCGAGGGGGGCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2282
Db CATGCTCTATCGAGGGGGGCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2285
QY 2283 CTCACATTCATTTGGCTTGGT 2342
Db CTCACATTCATTTGGCTTGGT 2345
QY 2343 CTACATGT 2402

QY 364 GCCACAGCCCGCTAGGAGTGTGGCCAGCCAGGTTGCTGTGGTCAAGCTTGCCACATC 423
Db |||||
QY 301 GCCCAGAGCCCGCTAGGAGTGTGGCCAGCCAGGTTGCTGTGGTCAAGCTTGCCACATC 360
Db |||||
QY 424 GAAGACTTCTCTCTGCAACCCCGAGTCCAGATTTGTGGAGGAGAAAGGACAGCAGCTTT 483
Db |||||
QY 361 GAAGACTTCTCTCTGCAACCCCGAGTCCAGATTTGTGGAGGAGAAAGGACAGCAGCTTT 420
Db |||||
QY 484 GAATGCGCACCAAGGGGCTTCCAGCCCGCATCATTTATTTGGGAAAAGGACAGGTGACC 543
Db |||||
QY 421 GAATGCGCACCAAGGGGCTTCCAGCCCGCATCATTTATTTGGGAAAAGGACAGGTGACC 480
Db |||||
QY 544 GTGCCCTGAGGAGCCCGGCTCATCTCTTCCCAAGTGGCTCTCCAGATCTTAGATGTC 603
Db |||||
QY 481 GTGCCCTGAGGAGCCCGGCTCATCTCTTCCCAAGTGGCTCTCCAGATCTTAGATGTC 540
Db |||||
QY 604 CAGGACAGTGATGCAAGGCTCTTACCGCTGCGTGGCCACCAATTCAGCCCGCCCAACGATTC 663
Db |||||
QY 541 CAGGACAGTGATGCAAGGCTCTTACCGCTGCGTGGCCACCAATTCAGCCCGCCCAACGATTC 600
Db |||||
QY 664 AGCCAGAGGCTCGTCTCATCTGTGGCCCTCAGAGGCTTTTGGAGGCTACCAAGGGGCGAG 723
Db |||||
QY 601 AGCCAGAGGCTCGTCTCATCTGTGGCCCTCAGAGGCTTTTGGAGGCTACCAAGGGGCGAG 660
Db |||||
QY 724 GATGTGTCATTTGTGGCAGCCCGCAGAGAACACCAAGTGTCTGTGGACAGATGTAGTG 783
Db |||||
QY 661 GATGTGTCATTTGTGGCAGCCCGCAGAGAACACCAAGTGTCTGTGGACAGATGTAGTG 720
Db |||||
QY 784 ATGAGTGTGCTGGCTCTGTGACCCCGCTTTTGTGCTGCTGGTCCGACAGGATGGA 843
Db |||||
QY 721 ATGAGTGTGCTGGCTCTGTGACCCCGCTTTTGTGCTGCTGGTCCGACAGGATGGA 780
Db |||||
QY 844 AAGCTATCTTCCAGGATGTATGCTGTGGCCCGGACCAATCTACTCATGCCAGCGCG 903
Db |||||
QY 781 AAGCTATCTTCCAGGATGTATGCTGTGGCCCGGACCAATCTACTCATGCCAGCGCG 840
Db |||||
QY 904 CAGCTTCGGCACTCTGAGTCTATGTCTGCGGAGCCAAAGCCCTCATCGGTGACTTC 963
Db |||||
QY 841 CAGCTTCGGCACTCTGAGTCTATGTCTGCGGAGCCAAAGCCCTCATCGGTGACTTC 900
Db |||||
QY 964 GCACTGCGGCTGTGAGTCTCGAGTGTCTGCTGCCCGCAGCCATCTCGCAGGCAACCCGAG 1023
Db |||||
QY 901 GCACTGCGGCTGTGAGTCTCGAGTGTCTGCTGCCCGCAGCCATCTCGCAGGCAACCCGAG 960
Db |||||
QY 1024 GCGTCTCGCGGACCGCGGCCAGCACCGCGGCTTGTGTGCGCGGCTCGCGGGAGCCA 1083
Db |||||
QY 961 GCGTCTCGCGGACCGCGGCCAGCACCGCGGCTTGTGTGCGCGGCTCGCGGGAGCCA 1020
Db |||||
QY 1084 CGGCCCGGCTGCACTGCGTGCACGACCGGATCCCGTTGGACCCCAATGGGCGGCTCAAG 1143
Db |||||
QY 1021 CGGCCCGGCTGCACTGCGTGCACGACCGGATCCCGTTGGACCCCAATGGGCGGCTCAAG 1080
Db |||||
QY 1144 GTGCAGGCGGTGGCGGAGCTTGTGTATCTCATGATCGGCTGCGAGGAGCTGGCTAC 1203
Db |||||
QY 1081 GTGCAGGCGGTGGCGGAGCTTGTGTATCTCATGATCGGCTGCGAGGAGCTGGCTAC 1140
Db |||||
QY 1204 TACCAAGTGTAGCAGAGAAAAGCGCGGGAACTGCTGTGCGCTGCGCCCTCGCGCGTA 1263
Db |||||
QY 1141 TACCAAGTGTAGCAGAGAAAAGCGCGGGAACTGCTGTGCGCTGCGCCCTCGCGCGTA 1200
Db |||||
QY 1264 GTGFTGCGGAGGGCTGCCCCAGCCCGGACTCGGGTCAAGCCACCGCCCTGAGCAGC 1323
Db |||||
QY 1201 GTGFTGCGGAGGGCTGCCCCAGCCCGGACTCGGGTCAAGCCACCGCCCTGAGCAGC 1260
Db |||||
QY 1324 TCCTGTGCTGTGGCTGTGGAGCGGCTGAGTTGCAAGGAGCAATCATTTGGCTTC 1383
Db |||||
QY 1261 TCCTGTGCTGTGGCTGTGGAGCGGCTGAGTTGCAAGGAGCAATCATTTGGCTTC 1320
Db |||||
QY 1384 TCTCTTCACTACAAAAGGAGGAGTGGACATGTGGAGTACAGTTTGGAGTAAAC 1443
Db |||||
QY 1321 TCTCTTCACTACAAAAGGAGGAGTGGACATGTGGAGTACAGTTTGGAGTAAAC 1380
Db |||||
QY 1444 AATGACACCAAGAGCTCGAGGTTCCGGACCTGGAAACCCCAACAGGATTTAGTTCTAC 1503

Db |||||
QY 1391 AATGACACCAAGAGCTGCAAGTTCGGACCTTGGAAACCCCAACAGGATTTAGTTCTAC 1440
Db |||||
QY 1504 GTGTGGGCTACTCTCCAGCTGGGGCCAGCCGCAACCTCCAGCCAGCCCTGTGTGATACA 1563
Db |||||
QY 1441 GTGTGGGCTACTCTCCAGCTGGGGCCAGCCGCAACCTCCAGCCAGCCCTGTGTGATACA 1500
Db |||||
QY 1564 CTGAGCGATGTCCCAGCGCAGCACCCAGCTTACCTTGTTCAGGCCCAACCCCTCGGAC 1623
Db |||||
QY 1501 CTGAGCGATGTCCCAGCGCAGCACCCAGCTTACCTTGTTCAGGCCCAACCCCTCGGAC 1560
Db |||||
QY 1624 ATCAGGGTGGCATGCTGCTGCTCCCTCCAGCTTCCAGCTTCCAGCTGCTGAGTAC 1683
Db |||||
QY 1561 ATCAGGGTGGCATGCTGCTGCTCCCTCCAGCTTCCAGCTTCCAGCTGCTGAGTAC 1620
Db |||||
QY 1684 AAGATAGATGATGCTGCTGGGGAAGGAGATCAGGTTTTCTCCAGCTGCTGGAAT 1743
Db |||||
QY 1621 AAGATAGATGATGCTGCTGGGGAAGGAGATCAGGTTTTCTCCAGCTGCTGGAAT 1680
Db |||||
QY 1744 GAGACACAACTTACGTTAAACTCTTCCAGCCAAACAAAGTGTACCGAGTCCGATTTCA 1803
Db |||||
QY 1681 GAGACACAACTTACGTTAAACTCTTCCAGCCAAACAAAGTGTACCGAGTCCGATTTCA 1740
Db |||||
QY 1804 GCTGGCACTGGGCTGCTGCTATGGAGTCCCTTCTCAGTGGATGCGACACAGCACCTGGT 1863
Db |||||
QY 1741 GCTGGCACTGGGCTGCTGCTATGGAGTCCCTTCTCAGTGGATGCGACACAGCACCTGGT 1800
Db |||||
QY 1864 GTGCACAAACAGAGCCATGTTCCCTTGGCCCTGCGAATTTGAAGTGGAGGCAAGATG 1923
Db |||||
QY 1801 GTGCACAAACAGAGCCATGTTCCCTTGGCCCTGCGAATTTGAAGTGGAGGCAAGATG 1860
Db |||||
QY 1924 GAGTCCCTGTGTGTATGGCAGCCGCCCTTCCAGCCCAACCAAGTGTACCGAGTCCGATTTCA 1983
Db |||||
QY 1861 GAGTCCCTGTGTGTATGGCAGCCGCCCTTCCAGCCCAACCAAGTGTACCGAGTCCGATTTCA 1920
Db |||||
QY 1984 CTCTACTGGGAGAGTGGGAAACAGAGGAGGAGCAGATGTGTGACCGCCGCCCGGCGGT 2043
Db |||||
QY 1921 CTCTACTGGGAGAGTGGGAAACAGAGGAGGAGCAGATGTGTGACCGCCGCCCGGCGGT 1980
Db |||||
QY 2044 CGTGGAGATCAAGCTTGGGACGCTCGGCGCTGCGGCTGGAAGAGAAAGTGAAGCAGTAT 2103
Db |||||
QY 1981 CGTGGAGATCAAGCTTGGGACGCTCGGCGCTGCGGCTGGAAGAGAAAGTGAAGCAGTAT 2040
Db |||||
QY 2104 GAACTGACCCAGTTAGTCCCTGGCAGCGCTGTAACGAGTGAAGCTCGTAGCTTTCAACAAA 2163
Db |||||
QY 2041 GAACTGACCCAGTTAGTCCCTGGCAGCGCTGTAACGAGTGAAGCTCGTAGCTTTCAACAAA 2100
Db |||||
QY 2164 CACAGGACGCTACGCTGCTGTGTGGAAGGCGAAGCAGGAGAGGCGCCCAACGCGAC 2223
Db |||||
QY 2101 CACAGGACGCTACGCTGCTGTGTGGAAGGCGAAGCAGGAGAGGCGCCCAACGCGAC 2160
Db |||||
QY 2224 CTGCTTATCCAGAGGGGCGCACCGCTGCTCTGCTCCCAATGTCACGAGAGTCAACAGC 2283
Db |||||
QY 2161 CTGCTTATCCAGAGGGGCGCACCGCTGCTCTGCTCCCAATGTCACGAGAGTCAACAGC 2220
Db |||||
QY 2284 TCCACTTTCCATTTGGCTTGGTGAAGAGCAGACTTTACCACTGTCAAGATTTGTCAAC 2343
Db |||||
QY 2221 TCCACTTTCCATTTGGCTTGGTGAAGAGCAGACTTTACCACTGTCAAGATTTGTCAAC 2280
Db |||||
QY 2344 TACACTGTACGCTTGGGCGCTGAGGAGTCTTCCCTGGTCACTACTATACC 2403
Db |||||
QY 2281 TACACTGTACGCTTGGGCGCTGAGGAGTCTTCCCTGGTCACTACTATACC 2340
Db |||||
QY 2404 AGCTCTGGAGAGACATTTCTCATTTGGCGGCTGAAACCAATTTACCAAGTACGAGTTTGG 2463
Db |||||
QY 2341 AGCTCTGGAGAGACATTTCTCATTTGGCGGCTGAAACCAATTTACCAAGTACGAGTTTGG 2400
Db |||||
QY 2464 GTACAGTCCCAACGAGTGGATATGATGGGCGCTTTTGGCTCGCTGAGAGAGCTTCCACC 2523
Db |||||
QY 2401 GTACAGTCCCAACGAGTGGATATGATGGGCGCTTTTGGCTCGCTGAGAGAGCTTCCACC 2460
Db |||||
QY 2524 CTGCTGACCGGCTTCAACACCTCTTCTGACCTGCGGCTGAGCCCTCGACACCAATCC 2583
Db |||||


```
Db 1273 GTGGTGGCGAGGGCTGCTCCAGCGCCCGCCAGCGGGTCACTGCTACGCCACTGAGCAGC 1332
Qy 1324 TCCTCTGTGCTGTGCTGGAGCGGCTGAGTTGCACAGCGAGCAATCAATGGCTTC 1383
Db 1333 TCGCTGTGTGTGGCTGGAGCGGCGCGAGATGCACAGCGAGCAGATCATGGCTTC 1392
Qy 1384 TCTCTTCACTACCAAAAGGGAAGGGAGGTGAGCAATGTGGAGTACAGATTTCAGATTAAC 1443
Db 1393 TCTCTCACTACCAAGAGCAGCGGGCATGGACAATGTGGAATACCAAGTTTGCAGTGAAC 1452
Qy 1444 AATGACACACAGAGCTGCAAGTTTCGGACCTGGGAACCAACACGAGTATCAGTTCTAC 1503
Db 1453 AAGCACACACAGAACTACAGTTTCGGACCTGGGAACCAACACAGTATCAGTTCTAC 1512
Qy 1504 GTGGTGGCTACTCCAGCTGGGGCCAGCCGAACTCCAGCCAGCCCTGGTGCATACA 1563
Db 1513 GTGGTGGCTACTCCAGCTGGGGCCAGCCGAACTCCAGCCAGCTGGTGCACACA 1572
Qy 1564 CTGGACGATGTCCCGAGCGAGCACCCGAGCTTACCTTGTCCAGCCCCAACCCCTCGGAC 1623
Db 1573 CTGGATGATGTAGGGCTCTGAATCTCGAGTGGC----- 1608
Qy 1624 ATCAGGGTGCATGGCTGCCCTCCAGCTGAGCAATGGACAGGTGCTGAAGTAC 1683
Db 1609 -----AGCTTGGGCTGAGCAATGGGAGGTGAGTAC 1644
Qy 1684 AAGATAGAGTACGGTTTGGGGAAGAAAGATCAGGTTTCTCCACGAGGTGCTGGAAT 1743
Db 1645 AAGATAGATACGGTTTGGGGAAGAAAGATCAGATTTCTTCTACTGAGGTGCGGGAAT 1704
Qy 1744 GAGACAACTTACGTTAACTCACTTCAGCCAAACAAAGTGTACCGAGTCCGGATTCA 1803
Db 1705 GAGACACAGCTTATGCTGAACCTCGCTTCAGCCAAACAAAGTGTATCGATACGGATTG 1764
Qy 1804 GCTGGACCTGGCGCTGGCTATGAGTTCCTTCTCAGTGAATGACGACAGCACTGCT 1863
Db 1765 GCTGGTACAGCGCGGCTTCGGGGCCCCCTCCAGTGGATGCATCAGAGGACGCCAGT 1824
Qy 1864 GTGCACAAACAGAGCCTATGCTCCCTTGGCCCTGAGAAATGGAAGTGAAGGCAAGATG 1923
Db 1825 ATGCACAAACAGAGCCTATGCTCCCTTGGCCCTGAGAGTTGAAGTGAAGGCAAGATG 1884
Qy 1924 GAGTCCCTGGTGTATGAGGAGCGCCCTTCAACCCAGCAATCTCTGGATACAA 1983
Db 1885 GAGTCCCTGGTGTATGAGGAGCGCCCTTCAACCCAGCAATCTCTGGTACAA 1944
Qy 1984 CTCTACTGGGAGAGTGGGAACAGAGGAGGAGGAGTGTGACCGCCCCCAGGGGT 2043
Db 1945 CTATATTGGCGGAGGTGGGGCTGAGGAGGAGGCAATGGCGATCGCTGCAGGGGCG 2004
Qy 2044 CGTGGAGATCAAGCTTGGGACGCTGGGGCCGCTGGGCTGAAGAAAGTGAAGCAGTAT 2103
Db 2005 CGTGGAGACAGCTTGGGATGTGGGGCTGTCCGGCTCAAGAAAGTGAAGCAGTAT 2064
Qy 2104 GAACTGACCACTTATGCTGGCAGCGCGTGTACGAGTGAAGTGTGAGTTTCAACAA 2163
Db 2065 GAGTGACCACTAGTGTCTGGCGGCTGTACGAGTGAAGTGTGAGTTTCAACAA 2124
Qy 2164 CAGGAGGAGGCTACGCTGTGTGGAAAGGGCAGAGCGGAGGAGGCGCCAGCGCAG 2221
Db 2125 CATGAGGATGGCTATGACAGTGTGGAAGGGCAGAGCGGAGGAGGCGCGCAGCAG 2182
```

RESULT 7

```
US-09-754-997A-5
; Sequence 5, Application US/09754997A
; Patent No. US20020102551A1
; GENERAL INFORMATION:
; APPLICANT: Salbaum, Michael J.
; TITLE OF INVENTION: NOPE Polypeptides, Encoding Nucleic
; TITLE OF INVENTION: Acids and Methods of Use
; FILE REFERENCE: P-NI 4552
```

```
; CURRENT APPLICATION NUMBER: US/09/754,997A
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: US 60/174,496
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: US 60/205,789
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 825
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(825)
US-09-754-997A-5
```

```
Query Match
Best Local Similarity 13.4%; Score 825; DB 10; Length 825;
Matches 825; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2932 CGACAAAGCTCCACAGGAAAGCCCTTCCCGGATTTGCTCTCAGGCAACCCAGGAAAC 2991
Db 1 CGACAAAGCTCCACAGGAAAGCCCTTCCCGGATTTGCTCTCAGGCAACCCAGGAAAC 60

Qy 2992 CCAGCGCTCTACACAAAGAGCTCGGCTTGGGCTCTCCAGTGTCCCTGTGCGCCATGAGTTG 3051
Db 61 CCAGCGCTCTACACAAAGAGCTCGGCTTGGGCTCTCCAGTGTCCCTGTGCGCCATGAGTTG 120

Qy 3052 GAGTCCCTCGTGCATCTCCGCTCCCGAGGATTTGGTTCCTCCCAACCCCTCAGATGTGAAGAC 3111
Db 121 GAGTCCCTCGTGCATCTCCGCTCCCGAGGATTTGGTTCCTCCCAACCCCTCAGATGTGAAGAC 180

Qy 3112 AAGGCTGAAGTACACAGAGCTTATGGGTGGCAGTGTTCAGAGTATGCGGGCCACTCCCAAG 3171
Db 181 AAGGCTGAAGTACACAGAGCTTATGGGTGGCAGTGTTCAGAGTATGCGGGCCACTCCCAAG 240

Qy 3172 AGAAAGATCTCTGGGCTCAGGCGGGGACAAACTGGGCGAGGCTCTCGGCGAGGCTGT 3231
Db 241 AGAAAGATCTCTGGGCTCAGGCGGGGACAAACTGGGCGAGGCTCTCGGCGAGGCTGT 300

Qy 3232 GAGCTGCCCCAGGGTAGTGGTCCAGGCGGGCTCTGACCCGCTGTCTGCTGCTCAGG 3291
Db 301 GAGCTGCCCCAGGGTAGTGGTCCAGGCGGGCTCTGACCCGCTGTCTGCTGCTCAGG 360

Qy 3292 GGAACCGGGCAGACACTGCTCTGCAAGCCCTGGTGTATGACGGATATAAGAGCAACGG 3351
Db 361 GGAACCGGGCAGACACTGCTCTGCAAGCCCTGGTGTATGACGGATATAAGAGCAACGG 420

Qy 3352 AGAAAGAGCCCTCCCGAGCTGCAAGGATCAGGTGGAAGCTGAGGTCAITTTCCATCC 3411
Db 421 AGAAAGAGCCCTCCCGAGCTGCAAGGATCAGGTGGAAGCTGAGGTCAITTTCCATCC 480

Qy 3412 GACTTCGGTGCATCAAAAGGATGCTGACCTCCACCTCCAAAGACTGAGGCGAGGAA 3471
Db 481 GACTTCGGTGCATCAAAAGGATGCTGACCTCCACCTCCAAAGACTGAGGCGAGGAA 540

Qy 3472 CCAGTGTGTCAGAGACTCTGCTTCCACCTCTGAGGCTGTGGATCTGTCTCAAGAGCA 3531
Db 541 CCAGTGTGTCAGAGACTCTGCTTCCACCTCTGAGGCTGTGGATCTGTCTCAAGAGCA 600

Qy 3532 GACTGGCTGGGCGAGGAGCTGGGAGGGTGCACCAACCAACAGTGGGCGCAGAGGCTC 3591
Db 601 GACTGGCTGGGCGAGGAGCTGGGAGGGTGCACCAACCAACAGTGGGCGCAGAGGCTC 660

Qy 3592 ACCTGCTTGCAGAGAGCAGAGCTGCTCTGCTCTGCTCTGCTCAGACTCCAGGCCAGCACT 3651
Db 661 ACCTGCTTGCAGAGAGCAGAGCTGCTCTGCTCTGCTCTGCTCAGACTCCAGGCCAGCACT 720

Qy 3652 GCTATAGAGGAGGCGCCCTGGGGAAGAGCTGCCAGGCCCAAGCCCTGTGTCTTACAGTC 3711
Db 721 GCTATAGAGGAGGCGCCCTGGGGAAGAGCTGCCAGGCCCAAGCCCTGTGTCTTACAGTC 780
```



```

QY 1974 TGGATACAACTCTACTGGGAGAGGTGGGAACAGAGGAGGAGGAGATGTGACCGCCC 2033
Db 1713 GGGCTACAAAGCTGTACTACAGGAAGAAGGCGCAGGAGAA-----1754
QY 2034 CCCAGGGGCTCGTGAGATCAAGCTTGGACCTGCGGCCCGCTGCGGCTGAAGAGAAAGT 2093
Db 1755 -----TGGGCCCAATTTCTTGGATACCAAGGA 1781
QY 2094 GAAGCAGTATGAATGACCCAGTTAGTCCCTGGCAGGCGCTACAGAGTGAAGCTGTGAC 2153
Db 1782 CCTACTCTATCTCTCAGTGGCTTAGACCCCAAGAAATAATCATGTGAGACTCTCGC 1841
QY 2154 TTTCAACAAACAGAGGAGCGCTAGCTGCTGTGTGTGGAAGGCGAAGAGGAGAGCGGCC 2213
Db 1842 TTACAACAACATAGACGATGGCTATCAGGCGAGATCAGACTGTGACAC---TCCAGGATG 1898
QY 2214 CACGCCAGACTGCTATCTACAGAGGGGCGCACCGCTGCTCTGCCCATGTCCAGCAGA 2273
Db 1899 CGTGTCTGTTGTGATCGATGGTCCCTCTCCACCAACCCGCCCATCTCTATGCGAA 1958
QY 2274 GTCAACAGCTCCACTTCCATTTGGCTTTCGGTGGGAAGCCAGACTTTTACCACTGTCAA 2333
Db 1959 GGTAAACACCTCATCTTCCATCTTCTGCACTTGGAGGAGGCTGCAATTCACCGCTGCACA 2018
QY 2334 GATTGTCAACTACACTGTACGTTGGGCCCTTGGGGGCTCAGGAATGCTTCCCTGGTCA 2393
Db 2019 AATCATTAACATACACCATTCGCTGTAATCTGTTGGCCTGCAGAAATGCTTCTTGGTTCT 2078
QY 2394 CTACTATACAGCTCTGAGAGAGACTTCTCATTTGGCGGCTGAAACCATTTTACCAAGTA 2453
Db 2079 GTACCTTCAAAATCAGAAATCTCAGATGTTGGTCAAGGCTAGAAACCAAAACCAATA 2138
QY 2454 CGAGTTTTCGGGTACAGTCCACGAGTGGATATGGATGGGCGCTTTGGCTCCGTCGTAGA 2513
Db 2139 CGAATTTGCGGTTGATATGATGTGATCAGCTTTCAGTCTTGGAGCCCTGTAGTCTA 2198
QY 2514 AGCTTCACCTGCTGACCGGCTTCAACACTCTTCTGACCTGCGGCTGAGCGCCCT 2573
Db 2199 CCAATTCATCTCTCAGAAAGCACACAGCAGGCGCCACCACTTGGAGTAAAGTGACATTAAT 2258
QY 2574 GACACCATCCACGTTGGTTACACTGGTGTGTCCTCCCGGAGGAGGAGGAGGAGTGT 2633
Db 2259 AGAGATGACACTGCGCTGTTCTTGGAAACCCCTGATGGCCCGAGAAACAGTTGTGAC 2318
QY 2634 GGAGTATCTAATTTCTACAGCAACACACACCCAGCGCCGAAACACAGTGGACACTGCT 2693
Db 2319 CCGCTATACTATCTTATATGATCTAGGAAGGCTTGGATTCAGAGAGTGGCAGGCTT 2378
QY 2694 CACCAAGAGGAAACATCTTCAAGTGCAGAGGTCAATGGCCTAGAGAGTGACACTCGGTA 2753
Db 2379 ACACCGTGAAGGGGCAATAACATGGCTTTCAGTAAACCTTGGTAGCAGGAATGTGTA 2438
QY 2754 TTTCTTCAAGATGGAGCGCCGACAGAGTGGGCGCTGGGCGCTTTTCCCGCTTGCAGGA 2813
Db 2439 CATTTGCAAGATATCTGATCCAAATGAGTGGGAGAGGAGGAGGAGGAGGAGGAGGAGG 2498
QY 2814 TGTGATTACTCTGCAAGAGACATCTCAGACTC 2846
Db 2499 GCTGGCAGTACTTCCAAAGGAACCTCTGAATC 2531

```

RESULT 9

```

US-10-231-353-1
; Sequence 1, Application US/10231353
; Publication No. US20030023064A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian

```

```

; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20030023064A1el Human Phosphatases and Polynucleotides En
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/10/231.353
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: US/09/877,730
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3210
; TYPE: DNA
; ORGANISM: homo sapiens
; US-10-231-353-1

```

```

Query Match      5.1%; Score 316.6; DB 14; Length 3210;
Best Local Similarity 47.4%; Pred. No. 1.4e-83;
Matches 1210; Conservative 1; Mismatches 1270; Indels 72; Gaps 6;

QY 300 AGACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTCTG 359
Db 279 ATACATCAGTGAAGTGAAGGCGAGGCGAGGAGAGCAGTCCGATGAAGGATTTTATCAGTG 338
QY 360 TCTGGCCACACAGCCCGCTAGGAGTGGCCAGCCAGGTTGCTGTGTGCTCAAGCTTGCAC 419
Db 339 CTTGGCAATGAACAAATATGAGCCATCTTAGTCAAAAAGCTCATCTTGCCTTATCAAC 398
QY 420 ACTCGAAGACTTCTCTCTGCACCCCGAGTCCAGATTTGTGAGGAGAAACGGGACGACG 479
Db 399 TATTTCTGCATTTGAAGTCCAGCCAAATTTCCACTGAGGTCCACGAGGTGGAGTTGCTCG 458
QY 480 CTTTGAATGCAACACCAAGGGCTTCCAGCCGCCCATCATTACTTGGGAAAGAGACCAAGT 539
Db 459 ATTGTCATGCAAGATTTTCATCCACCCTCTCTGCACTATAACATGGGAGTTCAATCGGAC 518
QY 540 GACCGTGCCTGAG---GAGCCCGGCTCATCATCTTCCCAAGTGGCTCTTCCAGATCCT 596
Db 519 AACTCTACCTATGACTATGACAGGATAACTGCCCTTACCAACAGGAGTATTGCAAGTCTA 578
QY 597 AGATGTCCAGCAGTGTGATGATGAGGCTCTTACCGCTGCGTGGCCACCAATTCAGCCCGCA 656
Db 579 TGATGTGAGCAAGGAGATTTCTGGAATTTATGTTGTATTGCTGCCACTGTAGGCCACCG 638
QY 657 ACGATTGAGCCAGAGCGCTCGCTCATCTGTGGCCCTCAGAGGCTTTTGGAGGCTACCAAG 716
Db 639 ACGTAAAGTATGAGGCGCTCGCTAACTGTGATTTCCAG-----CTAAGGAGTCAAAATC 692
QY 717 GGGGCGAGATGTGATGATGATGAGGCTCTTGGGCGCCCGAGAGAACACCAAGTGTGTGACAGAA 776
Db 693 CTTCCACACACACCAATTTATAGCAGGTCCACAGAAATATAACAAATCTCTTCTCATCAGAC 752
QY 777 TGTAGTATGAGTGTGCTGCTGCTGACCCCGCCCGCCCTTTTGTGTCTGGGTCGAGCA 836
Db 753 TGTAGTTTGGATGATGAGGCGCAGGAATCCCAACCAATCATTTCTTGGAGCGGCT 812
QY 837 GGAATGAAAGCTTAT---CTCCAGGATGTATCGTTCTGGGCGGAGCAATCTACTCAT 893
Db 813 TGATCACAAAATCCATGATGTCTTTAATACTCGGGTACTTGGAAATGTTATCTCATGAT 872
QY 894 GCGGAGCGGCGAGCTCGGCACTCTGAGTGTATGTCTGCGAGGCGCAACAGCCCTCAC 953
Db 873 ATCTGATGTGAGGCTACAAATGCTGGAGTATATGTTTGTGGGCGCACTACCCCTGGGAC 932
QY 954 GCGTGAATTCGCCACTCGCGGCTGAGCTCCGAGTGTCTGCGCCCGGAGCAATCTCGCA 1013
Db 933 ACGCACTTTTACAGTTCTATGGCACTTTAACTGTATAGTCTCTCTCTCATTTGTTGA 992
QY 1014 GGCAACCGGAGCGCTCTCGGCGAGCGCGGCGCAGCACCGCGCTTCTGTGTCCCGGGCTC 1073
Db 993 ATGGCCAGAAAGTTTAAACAGGCTCGAGCTGGCACTGCTCGATTTGTGTGTGAGGAGCA 1052
QY 1074 CGGGGAGCCAGCGGCGGCTGCACTGCTGACGAGGAGTCCCGTTGCGACCAATG 1133

```

Db 1053 AGGAATCCCTCTCCCAAGATGTATGGTTGAAAAATGGAAGGAGATACATTCCGAATGG 1112
Qy 1134 GCCTGTCAAGGTGCAGGGCGGTGGCGGAGCTTGGTTCATCACTCAAGTGGCTGCAGGA 1193
Db 1113 TAGAATTAATGATACAAAGT-----AAATTTGGTAATTAACCAAGATTATTTCTGAGAA 1166
Qy 1194 CGCTGCTACTACCACTGGTGTAGCAGAAACAGCGCGGAATGCTGCTGCGCTGCGCC 1253
Db 1167 TGATGCTATTTATCACTGATGGCTGAGAATAGCCAAAGATCTATTTTATCTAGAGCCAG 1226
Qy 1254 CTTGGCGGTAGTGGTGGGAGGGCTGCCAGCGCCCGCATCGGGTCAACAGCCAGCC 1313
Db 1227 ACTGACTGTAGTGTATGAGAAGACAGACCCAGTGTCTCCCTATAATGTATCATGCTGAAC 1286
Qy 1314 GCTGAGCAGCTCTCTGTGCTGGTGGCTGGGAGCGCTGAGTTGCCAGCGAGCAAT 1373
Db 1287 CATGTCAAGCTCAGCCATTTCTTTAGCTGGGAGGCGCACTTTATAATTCAGACAAAGT 1346
Qy 1374 CATTTGCTTCTCTCTTCACTACCAAAAGGCAAGGGAGTGGACAAATGTGGAGTACCAATT 1433
Db 1347 CATTTGCTTCTCTGTAACCTACATGAAGCAGAAGGTTTAAATAATGAAGAGTATCAAGT 1406
Qy 1434 TGCAGTAAACATGACACACAGAGCTGCAGGTTTCGGGACCTGGAAACCCAAACAGGATTA 1493
Db 1407 AGTCATCGGAATGACACAACTCATTTATTTATGATGACTTAGAGCCTGCCAGCAATTA 1466
Qy 1494 TGAGTTCTACGTGGTGGCTTCTCCAGCTGGGGCGAGCGCAACCTCCAGCGCCAGCCCT 1553
Db 1467 TACTTTCTACATTTGTAGCATATATGCAATGGGAGCCAGCAGATGTCTGACCATGTGAC 1526
Qy 1554 GGTGCATACATGAGCAGTGTCCCGAGCGAGCAGCCCGAGCTTACCTTTGCCAGCCCAA 1613
Db 1527 ACAGAACTCTAGAGGATGTTCCCTGAGAGCTCTCGAAATTTAGTTTGACAAGTCAAG 1586
Qy 1614 CCCTCGGACATCAGGGTGGCATGGCTGGCCCTGCGCTCCAGCCTGAGCAATGGACAGT 1673
Db 1587 TCCCACTGATTTCTCATCTCTGGCTGCCAATCCAGGCCAATATCGCGGGCGCAAGT 1646
Qy 1674 GCTGAAGTACAAGATAGATGACGGTTTGGGAGGAGAGATCAGGTTTCTCCACCGAGT 1733
Db 1647 GGTGCTGATCGCTGTCTTTCCGCTTAAGTACTGAGATTCAATCAAGTTCTGAGCT 1706
Qy 1734 GCCTGAAATAGACACAACTTACGTAACTCACTTCAGCCCAAAACAAAGTGTACCGAGT 1793
Db 1707 CCGGGGACCCAGCATGAGTACCTTTTGGAGGCGCTGAAACCTGACAGTGTCTACCTGGT 1766
Qy 1794 CCGAATTCAGTGGCACTGGCGCTGTATGAGTCCCTTCTCACTGATGATGACAGCAG 1853
Db 1767 TCGGATTTACTGCTGCCACCAAGAGTGGGCTGGGAGAGTATCATGATGAGACTTTCATAG 1826
Qy 1854 GACACCTGGTGTGCACAAACAGAGCCATGTTCCCTTTGCCCTGCAAGAAATTTGAAGTGAG 1913
Db 1827 GAGCCCAAGACTACAAGCGTGAAGCCCTTAAGTCTCCAGAGTTGATTTGGAGCTCT 1886
Qy 1914 GGCAAGATGAGTCCCTGGTGTGTATGCGAGCGCCGCCCTTCACCCCAACCCAGATCTC 1973
Db 1887 GAACTGTACCACATTTCTGTGAGTGGCAGCAAGATGTAGAGGACACAGCTGCTATTCA 1946
Qy 1974 TGGATTAACAACCTTACTTGGGGAGAGTGGGAAACAGAGAGGAGGAGAGTGTGACCGCCC 2033
Db 1947 GGGCTACAAGCTGTACTACAAGGAAGAGGGCGAGCAGGAA----- 1988
Qy 2034 CCAGGGGGTGTGGAGATCAAGCTTGGGAGCTGGGCGCGCTGGCTGAAGAGAAGT 2093
Db 1989 -----TGGGCGCATTTTCTTGGATACCAAGGA 2015
Qy 2094 GAAGCAGTATGAATGACCCAGTTAGTTCCTGGCAGCGCCGTACGAGTGAAGTCTGTAGC 2153
Db 2016 CCTACTCTATCTCAGTGGCTTAGACCCCAAGAAATATCATGTGAGACTCTCTGGC 2075
Qy 2154 TTTCAACAAACACGAGGACGGCTACCGTGTGTGTGGAAGGGCAAGCAGGAGAGGGCC 2213
Db 2076 TTACAACAAATAGACATGGCTATCAGGCGAGATCAGACTGTGAGCAC---TCCAGGATG 2132

Qy 2214 CACGCAGACCTGCCTATTCAGAGGGGGCCACCGCTGCTCTCTGCCATGTCCACGAGA 2273
Db 2133 CGTGTCTGTCGTGATGCAATGGTCCCTCCACCAACCCACCATCTCTATGCGAA 2192
Qy 2274 GTCAAAAGCTCCACTTCATTTGGCTTGGTGGAGGAGGAGGAGCTTTTACACATGTCAA 2333
Db 2193 GGCTAAACACCTCATCTTCCATCTTCTGCACTGGAGGAGGCTGCAATTCACCGCTGCACA 2252
Qy 2334 GATTGTCAACTACACTGCTTCGGCCCTCGCGGCTCAGGAATGCTTCCCTGGTGCAC 2393
Db 2253 AATCATTAACATACCAATCCGCTGTAATCTTGGCTGAGAAATGCTTCTTTGGTTCT 2312
Qy 2394 CTACTATACCACTCTGGAGAGACATTTCTCAITTTGGCGGCGCTGAAACCAATTTACCAAGTA 2453
Db 2313 GTACCTTCAAAACATCAGAAACCTCAGATGTTGGTTCAAGGTCTAGAAACCAACCAATA 2372
Qy 2454 CGAGTTTGGGTACAGTCCCAAGGAGTGTATGATGGGCGCTTTGGCTCGCTGCTAGA 2513
Db 2373 CGAATTTGCGGTTGATTTACATGATGAGATCAGCTTTCCAGTCTTGGAGCCCTGTAGTCTA 2432
Qy 2514 AGCTTCCACCTGCTGACCGGCTTCAACACCTCTTCTGACCTGCGCTGAGCCCT 2573
Db 2433 CATTTCTACTCTTCCAGAGGACAGGCGCCACAGTTGAGTGAAGTGAATTAAT 2492
Qy 2574 GACCAATCCACCGCTTGGTTTACACTGGTGTCCCCCGAGGCGCCAAATGGTGAATTTGT 2633
Db 2493 AGAGGATGACACTGCGCTGGTTTCTTGGAAACCCCTGATGGCCAGAAACAGTTGTGAC 2552
Qy 2634 GGAGTATCTAATTTCTTACAGCAACACACACCGCGCCGAAACACAGTGAAGTGAAGTCT 2693
Db 2553 CGCTATATCTATCTTATATGATCTAGGAAGGCTTGGATTTGAGGAGTGGCAGGTTCT 2612
Qy 2694 CACCAACAGGAGAAACATCTTTCAGTGCAGAGTCCATGCGCTTGTAGAAACTTGGTAGCAAGTGA 2753
Db 2613 ACACCGTAGAGGGGCAATAACATGCTTTTGTAGAAACTTGGTAGCAAGTGA 2672
Qy 2754 TTTCTTCAAGATGGAGGCGCCAGACAGAGTGGGGCTTGGGCGCTTTTCCCGCTTGCAGGA 2813
Db 2673 CATTTCAAGATATCTGATCCATGAGTGGGAGAGGACCTTTTCAAAATTTCTGTGA 2732
Qy 2814 TGTGATTAATCTGCAAGAGACATTTCTCAGACTC 2846
Db 2733 GCTGCGAGTACTTCCAAAGGAAACCTCTGAATC 2765

RESULT 10

US-10-231-353-17
; Sequence 17, Application US/10231353
; Publication No. US20030023064A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20030023064A1el Human Phosphatases and Polynucleotides E
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/10/231,353
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: US/09/877,730
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3219
; TYPE: DNA
; ORGANISM: homo sapiens

Query Match

5.1%; Score 316.6; DB 14; Length 3219;

		Best Local Similarity 47.4%; Pred. No. 1.4e-83;		Matches 1210; Conservative 1; Mismatches 1270; Indels 72; Gaps 6;	
QY	300	AGACAGCGATGATGAGGAGCTCTT	AGGATCTGGAAGGTCACTGAGGGCAGCTATTCCTG	359	
DB	45	ATACATCATGTGAGTGAAGGCGAGG	CGAGGCGAGGAGCGATCCGATGAAGGATTTATCAAGTG	104	
QY	360	TCGTGCCCCACAGCCGCTAGGAGTGG	TGGCCAGCCAGGTTGTGTGGTCAAGCTTGGCCAC	419	
DB	105	CTTGGCAATGAACAATATGGAGCCAT	CTTATAGTCANAAGCTCATCTTGGCTTATCAAC	164	
QY	420	ACTCGAAGACTTCTCTCTGACCCCG	AGTCCAGATTTGGAGGAGAAACGGACAGCAGC	479	
DB	165	TATTTCTGATTTGAATTCAGCCAA	TTTCCACTGAGTCCAGAGGTGGAGTTGTCTG	224	
QY	480	CTTTGAATGCCACACAAGGGCTTCC	AGCCCCCATCATTAATTGGGAAAGGACCAAGT	539	
DB	225	ATTTCATGCAAGATTTTCATCCCA	CCCTCTGCAGTATAACATGGAGTTCAATCGGAC	284	
QY	540	GACCGTGCCTGAG---GAGCCCGG	CTCATCTCTTCCCAAGTGCTCTCCAGATCCT	596	
DB	285	AACCTTACCTATGACTATGGACAG	GATAACTGCCCTACCAACAGGAGTATTCAGATCTA	344	
QY	597	AGATGTCCAGGACAGTGTACAGGCT	CTCTACCGCTGGCGCCACCAATTCAGCCCGCA	656	
DB	345	TGATGTCAAGCAAGGATCTTGAA	ATATCTGTGTATGTCTGCACATGTAGCCCAAG	404	
QY	657	ACGATTCAGCCAGGAGGCTCGCTCA	CTCTGTGCCCCCTCAGAGGCTCTTTGGAGGCTAC	716	
DB	405	ACGTAAAGTATGGAGGCTCGCTAA	CTTAAGTGTGATTCAG-----CTAAGGAGTCA	458	
QY	717	GGGGCAGGATGTGTCATTTGGGAG	CGCCCGCAGAGAACCAACGGTAGTGTCTGGACAA	776	
DB	459	CTTCCACACACACCAATATATAG	CAGGTCCACAGAACATAACAATCTCTTCATCAG	518	
QY	777	TGTAGTGTAGAGTGTGGCTCTGCT	GACCCCAACCCCTTTGTGTCTGGTCCGACA	836	
DB	519	TGTAGTTTGGAAATGATGGCCAC	AGGAATCCCAACCAATCATTTCTTGGAGCGCT	578	
QY	837	GGATGAAAGCCAT---CTCCACG	ATGTCACTGTTGGCCGAGCAATCTACTCAT	893	
DB	579	TGATCACAATCCATTTGATGTCT	TAACTCTGGGTACTTTGGAATGTATCTCATGT	638	
QY	894	CGCCAGCGGACGCTCGGCACTCT	GAGTCTATCTCTGCGAGCAACAAGCCCTCAC	953	
DB	639	ATCTGATGTCAAGCTACAAATG	CTGAGTATATGTTGTGCGGCGCACTACCCCTGG	698	
QY	954	GGTGACTTCGCCACTGCGGCTG	AGCTCCGAGTGTCTGTCGCCCGCAGCCATCTG	1013	
DB	699	ACGCAACTTTACGTTGCTATGG	CAACTTAACTGTATTTAGCTCTCTCTTCA	758	
QY	1014	GGCACCAGGCGCTCTCGGAGCG	CGGCGCAGCACCGCGCTTCTGTTGCGCGGCGTC	1073	
DB	759	ATGGCCAGAAAGTTTAAAGAGC	CTCTGAGCTGGCTCTGATTTGTGTGTCAGGCA	818	
QY	1074	CGGGAGCCACGCGCGGCTGCACT	GTGCTGCAAGCGGATTCGCTTGGACCCAAATGG	1133	
DB	819	AGGAATCCCTCTCCCAAGATG	TCATGTGTGAAATAAGAGGAAGATACATTCGA	878	
QY	1134	GGCGTCAAGTGCAGGCGGTGG	CGGCGAGCTTGTCTATCACTCAGATCGGCGCTG	1193	
DB	879	TAGAAATTAATATGTAACA	CAGT-----AAATGGTAAATTAACCAAGAT	932	
QY	1194	CGCTGGCTACTACCACTGCTAG	CAGAAACAGCGCGGGAATGCTGTGCGCTGCGCC	1253	
DB	933	TGATGCTATTTATCAGTGTCA	TGGCTGAGNATAGCCAGGATCTATTTTATCTAG	992	
QY	1254	CTGCGGTAGTGTGCGGAGGGCT	GTGCCAGCGCCCGGCTCGGGTCAAGCCACG	1313	
DB	993	ACTGACTGTAGTGTGAGAGAC	AGACCCAGTCTCCCTATATATGTACATGCTG	1052	
QY	1314	GCTGAGCAGCTCTCTGTGTGG	CGCTGGAGCGGCTGTGATGTCAGCGAGCAAT	1373	

DB	1053	CATGCTCAAGCTCAGCCATCTTT	TAGCTGGGAGGGCCACTTTATATT	CAGACAAAGT	1112
QY	1374	CATTGGCTTCTCTCTTCACTAC	CAAAAGGCAAGGGAGTGACAAAT	GTGGATACCAAGT	1433
DB	1113	CATTGCCCTATTCTGTACACTA	CTAGTAAGCAGAAGGTTTAAAT	TAATGAAGATATCAAGT	1172
QY	1434	TGCAGTAAACAATGACACCA	CAGAGCTCGAGTTGGGACCTG	GAACCAACACGGATTA	1493
DB	1173	AGTCATCGGAATGACCAACT	CATTATATTAATTGATGACT	TTAGAGCCCTGCACGCAATTA	1232
QY	1494	TGAGTTCTACCTGGTGGCTACT	CCAGCTGGGGCCAGCGCAACCT	CCAGCCAGCCCT	1553
DB	1233	TACTTTCTACATTTAGCATATA	TGCCNAATGGGAGCCAGCCAG	ATGTCAGACATGTGAC	1292
QY	1554	GGTGATACACTGGAACGATGT	CCCAAGCGCAGCAACCCAGCT	TACCTTGTCCAGCCCAA	1613
DB	1293	ACAGAACTACTAGAGGATGT	TCCCTGAGACCTCTCGAAAT	TAGTTTGACAAGTGGAG	1352
QY	1614	CCCCTCGGACATCAGGTGG	CGATGGCTGCCCTGCCCTCC	AGCTCAGCCTGAGCAATG	1673
DB	1353	TCCCACTGATATTTCTCATCT	CTGCTGCTGCAATCCAGCCAA	TATTCGGCGGGGCCAAGT	1412
QY	1674	GCTGAAGTACAAATAGAGTAC	GGTGGGGAAGAGATCAGGT	TTTTCTCCACCGAGGT	1733
DB	1413	GGTGTGTATCGCTTGTCTT	CCGCTTAAGTACTGAGAA	TTCAATCCAAAGTCTG	1472
QY	1734	GCCTGGAAATGAGACACAA	CTTACCTTTCAGCCAAACAA	AGTGTACCGAGT	1793
DB	1473	CCCGGGACCAACGATAGTAC	CTTTTGGAAAGGCTGAAAC	CTGACAGTGTCTACCTGT	1532
QY	1794	CGGATTTTCAGCTGGCACT	GGCTATGGAGTCCCTCTC	AGTGTGATGAGCAAG	1853
DB	1533	TCCGATTTACTGCTGCCA	CCAGAGTGGGGCTGGGAG	ATCATCAGTATGCAATG	1592
QY	1854	GACACCTGGTGTGCACAC	ACAGGCAATGTTTCCCTT	TCGCCCTGCAGAA	1913
DB	1593	GACGCCAAAGCTACAAGCT	GAAAGCCCTTAAGTCTC	CAGAGTTGCAATTTGG	1652
QY	1914	GGCAAGATGGAGTCCCT	TGTGTATGGCAGCGCCCT	CTCACCCACCCAGATCTC	1973
DB	1653	GAACTGTACCACTTTCT	GTGAGTGGCAGCAAGAT	GTAGAGGACACAGTGTCTAT	1712
QY	1974	TGGATACAACTCTACT	CTGCGGAGAGGTGGGAA	CAGAGAGGAGGAGATGTG	2033
DB	1713	GGGCTACAAAGTGTACTA	CAGAGGAAGGGCAGCAG	AGGAA-----	1754
QY	2034	CCCAGGGGTGCTGGAGAT	CAAGCTTGGGACGTGGG	CGCGCTGGAAGAAAGT	2093
DB	1755	-----	-----	-----	1781
QY	2094	GAAGCAGTATGAACCTG	ACCCAGTTAGTCCCTTGG	CAGCGCGGTACGAGGTGAAGCT	2153
DB	1782	CCTACTCTACTCTCT	AGTGGCTTAGACCCCAAG	AAATATCATGTGAGACTCTG	1841
QY	2154	TTTCAAACAAACAGGAC	CGCTACGCTGTGTGTG	GAAGGGCAGAGAGAGCGCC	2213
DB	1842	TTACACAAACATAGAC	ATGCTATCAGGCAGATC	CAGACTGTACAGCAC---	1898
QY	2214	CACGCCAGACTGCTAT	CCAGAGGGGCCACCGCT	GCTCTCTGCCCATGTCC	2273
DB	1899	CGTGTCTGTTGCTGAT	CGCATGGTCCCTCTCT	CCACCCACCCACTCTCT	1958
QY	2274	GTCAAAACAGCTCTCA	CTTTCCATTTGGCTTGG	TGGAAGAGGACAGACTTT	2333
DB	1959	GGCTAAACCTCATCT	CTTCATCTCTGACTG	GAAGGGCTGCATTTAC	2018
QY	2334	GATTGTCAACTACAT	GTATACGCTTCGGCCCT	TGGGGCTCAGGAATGCT	2393
DB	2019	AATCAATTAACCTAC	CACTCCGCTGTATCT	CTGTGGCTGCGAATGCT	2078
QY	2394	CTACTATACAGCTCT	GGAGAGACATTTCT	CAITTTGGCGGCTGAA	2453
DB	2079	GTACCTTCAACAT	CAGAAACTCATGT	TGGTTCAAGGTCTAG	2138

Db	1717	ACAGAAATACCTCTAGAGGATGTTCCCTCGAGACCTCTCTGAATAATTAGTTTGTGACAAGTCGAAG	1778
Qy	1614	CCCTCGGACATCAGGCTGGCATGGCTGCCCTGCCCTCCAGCTCAGCAATGACAGGT	1673
Db	1777	TCCCACTGATATTCTCATCTCTGGCTGCCAATCCAGCCAAATATATCGCGGGCCAGT	1836
Qy	1674	GCTGAAGTACAAGATAGAGTACGGTTTGGGGAAGGAAGATACAGGTTTCTCCACCGAGGT	1733
Db	1837	GGTGTCTGATCGCTTGTTCTTTCCGCTTAAGTACTCGAGAAATTCGAATCTTCAATCCAAAGTCTTGAGCT	1896
Qy	1734	GCTCGGAATGAGACACAACTTACGTTAAACTCATCTTCAGCCCAACAAAGTGTACCGAGT	1793
Db	1897	CCCGGACCAACGATGATGATACCTTTTGGAGGCTCGAAACCTGACAGTGTCTACCTGGT	1956
Qy	1794	CCGATTTACGTGGCACTCGCGCTGGCTATGAGTCCCTTCTCAGTGGATGAGACACAG	1853
Db	1957	TCGGATTACTGCTGCCACAGAGTGGGGCTGGAGAGTCAATCAGTATGGACTTTCACATAG	2016
Qy	1854	GACACCTGGTGTGCACAAACAGAGCCATGTTCCCTTTGGCCCCCTGCAGAAATGAAGGTGAG	1913
Db	2017	GAGCCCAAGCTACAAAGCGTGAAGCCCTTAAGTCTCCAGAGTTGCAATTTGAGGCTCT	2076
Qy	1914	GCCAAAGATGGATCCCTGGTGGTGTCAATGGCAGCCGCCCTTCCACCCACCCAGATCTC	1973
Db	2077	GAATGTACCAACCAATTTCTGTGAGGTGGCAGCAAGATGTAGAGACACAGCTGCTATTCA	2136
Qy	1974	TGGATACAAACTCTACTGGGAGAGGTGGCAACAGAGAGGAGGACAGATGTCGACCGCC	2033
Db	2137	GGCTACAACTGTACTACAAGNAGAAGGGCAGCAGAGAA	2178
Qy	2034	CCCAGGGGTCTGGAGATCAAGCTTGGGACGTGGGCCCGTGGGCCCTGAAGAAGAAAGT	2093
Db	2179	-----TGGGCCCATTTTCTTGATACCAAGGA	2205
Qy	2094	GAAGCAGTATGAATGACCCAGTTAGTTCCTGGCAGGCCGTACAGGTGAAGCTCGTAGC	2153
Db	2206	CCTACTATACTCTCATGTGGCTTAGACCCACAGAAAGAAATATCATGTGAGACTCTCTGGC	2265
Qy	2154	TTTTCAACAAACAGAGGACGGCTACGCTGTGTGTGAAGGGCAAGACGGGAAGGCGCC	2213
Db	2266	TTACAAACATAGACGATGGCTATCAGGACGATCAGACTGTACAGCAC-----TCCAGGATG	2322
Qy	2214	CAGCCAGACCTTGCCTATCCAGAGGGGGCCACCGCTGCCTCTGCCCATGTCCACGCAGA	2273
Db	2323	CGTGTCTGTCGTATCGGATCGATGTCCTCTCCACCAACCCACCATCTCTATGCGAA	2382
Qy	2274	GTCAACAGCTCCATTCCTCATTTTGGCTTGGGTGGGAAGACGACGATTTTACACTGTCAA	2333
Db	2383	GGCTAAACCTCATCTTCCATCTTCTGTCACTGGAGGAGGCGTGCATTCACCGCTGCACA	2442
Qy	2334	GATTTCAACTACACTGTACGCTTCGGCCCTCGGGGCTCAGGAATGCTTCCCTGGTGCAC	2393
Db	2443	AATCAITTAACCTACACCATCCGCTGTAAATCTGTGTGGCTCTGAGAAATGCTTCTTGGTTCT	2502
Qy	2394	CTACTATACCAGCTCTGGAGAAGACATTTCTCATTTGGCGGCTCGAAACCATTTACCAAGTA	2453
Db	2503	GTACCTTCAAAATCAGAACTCACAATGTTGGTTCAAGTCTAGAACCAACCAACAAATA	2562
Qy	2454	CGAGTTTGGGTACAGTCCACGAGGTGGATATGATGGGCGCTTTTGGCTCCGTCGTAGA	2513
Db	2563	CGAATTTGGCGTTTCAATTAATGTGGATCAGCTTTCCAGTCTTGGAGCCCTGTAGTCTA	2622
Qy	2514	ACGCTCACCTTGCCTGACCGGCCCTTCAACCTCTTCTGACCTGGCGCTGAGCCCCCT	2573
Db	2623	CCATTTCTACTCTTCCAGAAGCAACAGCAGGCCCAACCAAGTTTGGAGTAAAGTGACATTAAT	2682
Qy	2574	GACACCATCAACCGTTCGGTTTACACTGGTGTCCCCCAACGAGGCCCAATGGTGAGATTGT	2633
Db	2683	AGAGGATGACACTGCCCTGGTTTCTTGGAAACCCCTGTATGGCCAGAAACAGTTGTGAC	2742
Qy	2634	GGAGTATTAATTCTCTACAGCAACCAACCAACCGGCCCAACCAAGTGGACACTGCT	2693
Db	2743	CCGCTAFACTATCTTATATGCTATCAGAAAGGCTGTGATTTGCAGAGAGTGGCAGGTCCT	2802

```

QY 2694 CACACAGAGGGGAAACATCTTTCAGTGCAGAGGTCCATGGCGCTTAGAGAGTGACACTCGGTA 2753
Db 2803 ACACCGTGAAGGGGCAATAACCATGGCTTTGCTAGAGAAACTTGGTAGCAGGAAATGTGTA 2862
QY 2754 TTTCTTCAAGTGGGAGCCGCGACAGAGGTGGGGCCCTTGGGCCCTTTTCCCGCTTGACGGA 2813
Db 2863 CATGTCTCAAGATATCTGTCATCCATGAGTGGGAGAGGACCCCTTTTCAAAATTTCTGTGGA 2922
QY 2814 TGTGATTACTCTGCAAGAGACATTTCTCAGACTC 2846
Db 2923 GCTGGCAGTACTTCCAAAGGAAACCTCTGAATC 2955

RESULT 12
US-10-231-353-7
; Sequence 7, Application US/10231353
; Publication No. US20030023064A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: NO. US20030023064A1el Human Phosphatases and Polynucleot
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/10/231,353
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: US/09/877,730
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 3453
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-231-353-7

```

Query Match	5.1%; Score 315.4; DB 14; Length 3453;
Best Local Similarity	47.4%; Pred. No. 3.4e-83;
Matches 1210; Conservative	0; Mismatches 1271; Indels 72; Gaps 6;
QY	300 AGACAGCGATGATGACGAAAGCTCTTAGATCTCGNAGGTCACTGAGGCGAGCTATTTCCTG 359
DB	
QY	279 ATACATCAGTGAGGTGAAGCGAGCGGCGAGAGAGCAGTCCGATGAGGATTTTATCAGTG 338
DB	
QY	360 TCTGGCCCA CAGGCCGCTAGGAGTGTGGCCAGCGAGTTGCTGTGTGTCAAGCTTGGCCAC 419
DB	
QY	339 CTTGGCAATGAACAAATATGAGACCAATCTTATAGTCAAAAAGCTCATCTTGCCCTTATCAAC 398
DB	
QY	420 ACTCGNAGACTTCTCTCTGCAACC CGAGTCCCAGATTGTGGAGGAGAAACGGACAGCACG 479
DB	
QY	399 TATTTCTGCAATTTGAGTCCAGCCCAATTTTCACTGAGGTCCA CGAAGTGAGATTGCTCG 458
DB	
QY	480 CTTTGAATGCCACACCAAGGCGCTTCAGGCCCCCATCATTACTTTGGGAAAGGACCAAGT 539
DB	
QY	459 ATTTGCAATCAAGATTTTCATCCACCCCTCTCGAGTCAATAACATGGAGTTCAATCGGAC 518
DB	
QY	540 GACGTCGCTGAG---GAGCGCGGCTCATCATCTTCCCAAGTGGGCTCTCTCAGATCCT 596
DB	
QY	519 AACTCTACCTATGACTATGACAGGATAA CTGCGCTTACCAACAGGAGTATTCGAGATCTA 578
DB	
QY	597 AGATGTCCAGGACAGTATGACAGGCTCTTACCGCTGCTGGGCCACCACTTTCAGCGCCGCCA 656
DB	
QY	579 TGATGTGAGCCAAAGGATTTCTGGAAATATGTTGTATTGCTGCCACTGTAGGCCACCG 638
DB	
QY	657 ACGATTGAGCCAGAGGCGCTCGCTCATCTGTGGCCCTCAGAGGCTCTTTGAGAGGTATCACG 716
DB	
QY	639 ACGTAAAGATGTGAGAGCCCTCGCTTAATCTGATTTCCAG-----CTAAGGAGTCAAAATC 692
DB	
QY	717 GGGGCGAGATGTGCTCAATTGTGGCAGCGCCACAGAGAAACCA CCGTGTGCTGTGGACAGAA 776
DB	

Db 693 CTTCCACACACCAACAAATATAGCAGGTCCACAGAAATAAACAAATCTCTCATCAGAC 752
Qy 777 TGTAGTATGAGTGGTGGCTCTGCTGACCCCAACCCCTTTTGTGCTGCTGGGTCCGACA 836
Db 753 TGTAGTTTGGAAATGCAATGGCCACAGAAATCCCAACCAATCAATTTCTTGGAGCCGCT 812
Qy 837 GGATGGAAGCCAT---CTCCACGGATGTATCGTTCTGGGCGGACCAATCTACTCAT 893
Db 813 TGATCACAAATCCATTCATGCTTTTAATATCTCGGGTACTTGGAAATGGTAAATCTCATGAT 872
Qy 894 CGCAGGCGCAGGCTCGGCACTCTGAGTCTATGCTGCGAGCCCAACAGCCCTCAG 953
Db 873 ATCTGATGTGAGGCTACAAATGCTGAGTATATGTTTGTGCGGCCACTACCCCTGCGAC 932
Qy 954 GCGTGACTTCGCCACTCGGCTGCTGAGCTCCGAGTCTTGTGCCCCAGCCATCTCGCA 1013
Db 933 ACGCACTTTTACAGTTGCTATGGCACTTTAACTGTTATTAAGTCTCTCTCATTTGTTGA 992
Qy 1014 GGCAACCGAGCGCTCTCGGGAAGCGGGCCAGCACCGCGCTCTGTTGTCGGGCGTC 1073
Db 993 ATGSCCAAGAAATTTAAACAGGCTCGAGCTGGCACTGCTCGATTTGTGTGTCAGGCAGA 1052
Qy 1074 CGGGAGCCAGGCGCGCTGCACTGCTGCTGCAAGCGGATCCGTTGGCAACCAATGG 1133
Db 1053 AGGAATCCCTCTCCCAAGATGTATGTTTGAATAATGGAAGGAAGATACATTCGAATGG 1112
Qy 1134 GCGGTCAAGTGCAGGCGGCTGGCGGAGCTTGGTCACTCAGATCGGCTCGAGGA 1193
Db 1113 TAGAATTAATGTAACACAGT-----AAATTGGTAAATTAACAGATTAATCTCTGAAGA 1166
Qy 1194 CGGTGGCTACTACAGTGCCTAGCAGAAACAGCGCGGAATCGCTGTGCGCTGGCC 1253
Db 1167 TGATGCTATTTATCAGTGCATGCTGAGATAGCAAGGATCTATTTATCTAGAGCCAG 1226
Qy 1254 CTTGGCGGTAGTGTGCGAGGCGCTGCCAGGCGCCGACTCGGGTCAAGCCAGCC 1313
Db 1227 ACTGACTGATGATGTGAGAAAGACAGACCCAGTGTCTCTTAATATGATGATGAGGCTGCGAAT 1286
Qy 1314 GCTGAGAGCTCTCTGCTGTGCTGGCTGGAGCGGCTGAGTTGACAGCGGCAAT 1373
Db 1287 CATGTCAAGTTCAGCCATCTTTTGGCTGGAGAGGCCACTTTATTAATTCAGACAAAT 1346
Qy 1374 CATTTGGCTCTCTCTTCACTACCAAAAGGCAAGGGAGTGGCAATGTGGAGTACCAAGTT 1433
Db 1347 CATTTGGCTATTTCTGTACACTACATGAAGCAGAGGTTTAAATTAATGAAGATCAAGT 1406
Qy 1434 TGCAGTAAACAAATGACACCAAGAGCTGCGAGTTTCGGGACCTGGAAACCAACAGGATTA 1493
Db 1407 AGTCATCGGAAATGACACAACTCAATTAATTAATGATGACTTAGAGCTGCGAGCAATTA 1466
Qy 1494 TGAGTTCTAGTGGTGGCTACTCCAGCTGGGGCCAGCGCAACCTCCAGCGCCCT 1553
Db 1467 TACTTTCTACATTTGATGATATATGCAATGGGAGCCAGCCAGATGTCTGACCATGTGAC 1526
Qy 1554 GGTGCATACACTGACCAATGTCCCGAGCGAGCACCCAGCTTACCTTGTCCAGCCCA 1613
Db 1527 ACAGAAATCTCTAGAGATGTTCCCTGAGACCTCTCGAAATAGTTTGAACAGTGAAG 1586
Qy 1614 CCCCCTGGACATCAGGGTGGCAATGGCTGCCCTCCGCTCCAGCGCTGAGCAATGACAGGT 1673
Db 1587 TCCCACTGATATCTCACTCTCGCTGCCAATCCAGGCCAATATCGCGGGGCCAAGT 1646
Qy 1674 GCTGAAGTACAGATAGAGTACGGTTTGGGGAAGAGATCAGGTTTTTCTCCACCGAGT 1733
Db 1647 GGTGCTGATGCTGTGTTCTTCGCGCTTAAGTACTGAGAAATTCATTCGAATTCGAGCT 1706
Qy 1734 GCCTGGAAATGAGACAACTTACGTTAAATCACTTTCAGCAAAACAAAGTGTACCGAGT 1793
Db 1707 CCGGGGACCAACGATGATACCTTTTGGAGGCTGAAACCTGACAGTGTCTACCTGGT 1766
Qy 1794 CCGGATTTCACTGCGCTGCGCTAGGCTATGAGTCCCTTCTCAGTGGATGACAGCAG 1853

Db 1767 TCGGATTAATCTGCTGCCACACAGAGTGGGCTGGGAGAGTCAATCAGTATGACTTTCACATAG 1826
Qy 1854 GACACCTGTGTGCAACACAGAGCATGTTCCCTTTTGCCTTGCAGAAATTCAGAGTGTAG 1913
Db 1827 GACGCCAAAGCTACAAGGTGAAGCCCTTAAGTCTCCAGAGTTGCAATTTGGAGCCTCT 1886
Qy 1914 GGAAGAATGAGTATCCCTGCTGCTGTGTGTATGAGCGGCGCCCTCACCACCCAGATCTC 1973
Db 1887 GAACTGTACCACTTTCTGTGAGGTGGCAGCAAGATGTAGAGGACACAGCTGCTATTCA 1946
Qy 1974 TGGATACAACTCTACTGGGAGAGTGGGAACAGAGGAGGAGGAGATGGTGACCGCC 2033
Db 1947 GGGCTACAAGCTGTACTACAAGAAAGAGGCGAGGAGAA----- 1988
Qy 2034 CCCAGGGGTCTGGAGATCAAGCTTTGGGACGTCCGCGCCCTGGCGCTGGAAGAGAAAGT 2093
Db 1989 -----TGGGCCCATTTTCTTGGATACCAAGA 2015
Qy 2094 GAAAGATATGAATGACCCAGTTAGTCTCTGGAGGCGGTACAGAGTGAAGCTGTAGC 2153
Db 2016 CTTACTCTATCTCTCAGTGGCTTAGACCCAGAAAGAAATATCATGTGAGACTCTCTGCG 2075
Qy 2154 TTTCACAAACACAGGAGCGCTACGCTGTGTGTGGAAGGCAAGACGAGAGGCGCC 2213
Db 2076 TTACAAACACATAGACGATGGCTATCAGGCAGATCAGACTGTGAGC-----TCCAGGATG 2132
Qy 2214 CACGCCAGACCTGCGCTATCCAGAGGGGCCACCGCTGCTCTGCGCCATGTCCACGACAG 2273
Db 2133 CGTGTCTGTGTGATCGATGGTCCCTCTCCACCAACCCACCATCTCTATGCGAA 2192
Qy 2274 GTCAAAACAGCTCACTTTCATTTTGGCTTGGTGAAGAAAGACAGACTTTTACACTGTGCA 2333
Db 2193 GGTAAACACTCATCTTCCATCTTCTGCACTGGAGGAGCTTGCATTCACGCTGACAC 2252
Qy 2334 GATTGTCAACTACATGTAGCTTGGCGCTTGGGGCTCAGGAAATGCTTCTCTGCTGAC 2393
Db 2253 AATCATTAATACACATCCGCTGTAATCTTGTGGCTTGCAGAAATGCTTCTTGTGTTCT 2312
Qy 2394 CTACTATACAGCTCTGGAGAAAGACATTTCTCATTTGGCGCTGAAACCAATTTACCAAGTA 2453
Db 2313 GTACCTTCAAAATCAGAAACTCACAATGTTGGTTCAGGTCTTAGAACCAACCAATA 2372
Qy 2454 CGAGTTTGGGTACAGTCCCAAGAGTGTATGAGATGGGCTTTTGGCTCGGTCTGATA 2513
Db 2373 CGAAATTTGCGTTGATTTACATGTGATCAGCTTTTCCAGTCTTGGAGGCTGTAGTCTA 2432
Qy 2514 AGCTTCACTGCTGCTGACCGGCTTCAACACTCTTCTGACCTGCGCTGAGCCCT 2573
Db 2433 CCATTTCTCTTCCAGAAAGCACAGAGGCCACCAAGTTGGAGTAAAGTGAATTAAT 2492
Qy 2574 GACACCATCCCGTTTCGGTTACACTGTGTTCCTCCCGACGGAGCCCAATGGTGAGATTGT 2633
Db 2493 AGAGGATGACATGCTGCTGTTTCTTGGAAACCCCTGTATGGCCAGAAACAGTTGTGAC 2552
Qy 2634 GAGATATCTAATTTCTTACAGCAACAAACACACCCAGCCGAAACACAGTGGACATGCT 2693
Db 2553 CGCTATATCTATCTTATATGCTCTAGGAAGGCTGGAATTCAGAGAGATGGCAGTCTT 2612
Qy 2694 CACCACAGAGGAAACATCTTCACTGACAGGCTTCATGGCTTAGAGGCTAGAGTGAACACT 2753
Db 2613 ACACCGTGAAGGGCAATTAACATGGCTTTTGTAGAAATCTTGGTAGCAGGAATGTGA 2672
Qy 2754 TTTCTTCAAGATGGGAGCCCGACAGAGTGGGCTTGGCGCTTTCGCTTTTCCGCTTGCAGGA 2813
Db 2673 CATTTGCAAGATATCTGATCCAAATGAGTGGGAGAGAGACCTTTTCAATTTCTGTGA 2732
Qy 2814 TGTGATTACTCTGCAAGAGACATTTCTCAGATC 2846
Db 2733 GCTGGCAGTACTTCCAAAGGAACCTCTGAATC 2765

; Sequence 19, Application US/09754997A
; Patent No. US20020102551A1
; GENERAL INFORMATION:
; APPLICANT: Salbaum, Michael J.
; TITLE OF INVENTION: NOPE Polypeptides, Encoding Nucleic
; TITLE OF INVENTION: Acids and Methods of Use
; FILE REFERENCE: P-NI 4552
; CURRENT APPLICATION NUMBER: US/09/754,997A
; CURRENT FILING DATE: 2001-01-04
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: US 60/174,496
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: US 60/205,789
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 288
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(288)
US-09-754-997A-19

Query Match 4.7%; Score 288; DB 10; Length 288;
Best Local Similarity 100.0%; Pred. No. 1e-75;
Matches 288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1888 TTGTGCCCCGCGAGATTGAAGTGGAGGCGCAAGATGGAGTCCCTGGTGTCTCATGGCAG 1947
Db 1 TTGTGCCCCGCGAGATTGAAGTGGAGGCGCAAGATGGAGTCCCTGGTGTCTCATGGCAG 60
QY 1948 CGGCCCCCTCACCCACCCAGATCTCTGGATCAAACTCTACTGGGGAGAGGTGGGAACA 2007
Db 61 CGGCCCCCTCACCCACCCAGATCTCTGGATCAAACTCTACTGGGGAGAGGTGGGAACA 120
QY 2008 GAGGAGGAGCGAGATGGTGACCGCCCCCGCGGGGTCTGGAGATCAAGCTTTGGACGTC 2067
Db 121 GAGGAGGAGCGAGATGGTGACCGCCCCCGCGGGGTCTGGAGATCAAGCTTTGGACGTC 180
QY 2068 GGGCCGCGCGCTGAAGAGAGAGTGAAGCAGTATGAAGTCAAGCTTTAGTCCCTGGC 2127
Db 181 GGGCCGCGCGCTGAAGAGAGAGTGAAGCAGTATGAAGTCAAGCTTTAGTCCCTGGC 240
QY 2128 AGGCGGTACGAGGTGAAGTCTGTAGCTTTCAACAAACACGAGGAGCGC 2175
Db 241 AGGCGGTACGAGGTGAAGTCTGTAGCTTTCAACAAACACGAGGAGCGC 288

RESULT 14
US-09-754-997A-23
; Sequence 23, Application US/09754997A
; Patent No. US20020102551A1
; GENERAL INFORMATION:
; APPLICANT: Salbaum, Michael J.
; TITLE OF INVENTION: NOPE Polypeptides, Encoding Nucleic
; TITLE OF INVENTION: Acids and Methods of Use
; FILE REFERENCE: P-NI 4552
; CURRENT APPLICATION NUMBER: US/09/754,997A
; CURRENT FILING DATE: 2001-01-04
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: US 60/174,496
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: US 60/205,789
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 252
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(252)

US-09-754-997A-23

Query Match 4.1%; Score 252; DB 10; Length 252;
Best Local Similarity 100.0%; Pred. No. 6.5e-65;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2542 ACACCTCTCTTGACCTGGCTGAGCCCTGAGCCCTGACACCATCCACCGTTTCGTTTACACTGG 2601
Db 1 ACACCTCTCTTGACCTGGCTGAGCCCTGAGCCCTGAGCCCTGACACCATCCACCGTTTCGTTTACACTGG 60
QY 2602 TGTCCCCCAGGAGCCCAATGGTGAGATTCTGGAGTATCTAATTTCTTACAGCAACAC 2661
Db 61 TGTCCCCCAGGAGCCCAATGGTGAGATTCTGGAGTATCTAATTTCTTACAGCAACAC 120
QY 2662 CACACCCAGCCCGAACACCACTGCTCACACAGAGGGAACATCTTCAAGTCA 2721
Db 121 CACACCCAGCCCGAACACCACTGCTCACACAGAGGGAACATCTTCAAGTCA 180
QY 2722 GAGGTCCATGGCCCTAGAGAGTGACACTGCGGTATTTCTTCAAGTGGAGCCCGCACAGAG 2781
Db 181 GAGGTCCATGGCCCTAGAGAGTGACACTGCGGTATTTCTTCAAGTGGAGCCCGCACAGAG 240
QY 2782 GTGGGGCTGGG 2793
Db 241 GTGGGGCTGGG 252

RESULT 15
US-09-754-997A-15
; Sequence 15, Application US/09754997A
; Patent No. US20020102551A1
; GENERAL INFORMATION:
; APPLICANT: Salbaum, Michael J.
; TITLE OF INVENTION: NOPE Polypeptides, Encoding Nucleic
; TITLE OF INVENTION: Acids and Methods of Use
; FILE REFERENCE: P-NI 4552
; CURRENT APPLICATION NUMBER: US/09/754,997A
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: US 60/174,496
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: US 60/205,789
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 249
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(249)
US-09-754-997A-15

Query Match 4.0%; Score 249; DB 10; Length 249;
Best Local Similarity 100.0%; Pred. No. 5.2e-64;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1285 AGCGCCCGACTCGGTCAGAGCCAGCCGCTGAGCAGCTCTCTGTGCTGGTGGCCTGG 1344
Db 1 AGCGCCCGACTCGGTCAGAGCCAGCCGCTGAGCAGCTCTCTGTGCTGGTGGCCTGG 60
QY 1345 GAGCGGCTGAGTTGCAAGCAATCAATGGCTTCTCTTCACTACCAAAAGCA 1404
Db 61 GAGCGGCTGAGTTGCAAGCAATCAATGGCTTCTCTTCACTACCAAAAGCA 120
QY 1405 AGGGAGTGGCAATGTGGAGTACCAGTTTGAGTAAACATGACACCAAGAGCTGAG 1464
Db 121 AGGGAGTGGCAATGTGGAGTACCAGTTTGAGTAAACATGACACCAAGAGCTGAG 180
QY 1465 GTTCGGGACCTGGAAACCCAAACACCGATTAAGATTCTACGTGGTGGCCTACTCCAGCTG 1524
Db 181 GTTCGGGACCTGGAAACCCAAACACCGATTAAGATTCTACGTGGTGGCCTACTCCAGCTG 240

Qy 1525 GGGGCCAGC 1533
| | | | |
Db 241 GGGGCCAGC 249

Search completed: October 15, 2003, 12:28:06
Job time : 1679.34 secs

Qy	597	AGATGTCAGGAGTGAATGAGGCTCTTACCGCTGCTGGCCACCAATTCAGCCGCCA	656	1413	GGTGTGTATCGCTTGTCTTCCGCTAAGTACTGAGAAATCAATCAAGTCTTGAGCT	1472
Db	345	TGATGTGAGCCAAAGGATTCGGAATATATGTTGTAATGCTGCCACTGTAGCCACCG	404	Qy	1734	GCTGGAAATGAGACACAACTTACGTTAACTCACTTCAGCCAAACAAAGTGTACCGAT
Qy	657	ACGATTCAGCAGGAGGCTCGCTCACTGTGGCCCTCAGAGGCTCTTGGAGGCTACCG	716	Db	1473	CCGGGACCAAGCATGAGTACTTCTTGGAGGCTGAAACCTGACAGTGTCTACTCTGT
Db	405	ACGTAAGATGATGAGGCTCGCTAATCTGTGATTCAG-----CTAAGGATCAAAATC	458	Qy	1794	CCGATTTTCACTGAGCTGGGCTGCTGAGTGTCTTCTCAGTGTGATGACGACAG
Qy	717	GGGACAGATGCTGCTATTTGGCAGGCCACAGAGAACACACGCTAGTGTCTGACAGAA	776	Db	1533	TGGATTAATCTGCTGCCACAGAGTGGGCTGGGAGAGTCAATGATGAGCTTCAATAG
Db	459	CTTCCACACACCAATATAGCAGGTCCACAGAACATTAACACATCTCTTCATCAGAC	518	Qy	1854	GCACCTGTGTGTCACAAACAGGACCATGTTCCCTTTGGCCCTTGCAGAAATGAGGTGAG
Qy	777	TGTAGTGAATGAGTGTGGCTCTGCTGACCCCAACCTTTTGTGTGCTGGGTCCGACA	836	Db	1593	GACGCCAAAGCTTACAAGCGTGAAGCCCTCAAGTCTCCAGAGTTGCTATTGGAGCTCT
Db	519	TGTAGTTTGGAAATGATGAGGCTCGCTAATCTGTGATTCAG-----CTAAGGATCAAAATC	578	Qy	1914	GGCAAGATGAGTGTCTGCTGTCATGTCAGGCGCCGCTCCCTCAACCCACAGATCTC
Qy	837	GGATGGAAGCTAT---CTTCCAGGATGTCTGTTCTGGCCGAGCAATCTACTCAT	893	Db	1653	GAACTGTACCAACATTTCTGTGAGTGGGAGCAAGATGTAGAGGACAGCTGCTATTCA
Db	579	TGATCACAAAATCCATGATGTCTTAAATCTCGGGTACTTGGAAATGGTAAATCTCATGAT	638	Qy	1974	TGGATACAAATCTTACTTGGGAGAGGTGGGAAACAGAGGAGGAGGAGATGGTGACGCGCC
Qy	894	CGCCAGGCGCAGGCTCGGCACTCTGGAGTCTATGTCTGCGAGCCAAACAGCCCTCAC	953	Db	1713	GGGCTACAAGCTGTACTACAGGAAGAGGCGAGCAGAGAA-----
Db	639	ATCTGATGTAGGCTCAACATGCTGGAGTATATGTTGTGCGGCCACTACCCCTGGCAC	698	Qy	2034	CCAGGGGCTGTGGAGATCAAGCTTGGGACGTGGGCGCGCTGGGCTGAAGAGAAAGT
Qy	954	CGGTGACTTCCCACTGCGGCTGTGAGCTCCGAGTGTGCTGCCCGAGCAATCTCGCA	1013	Db	1755	-----TGGGCGCAATTTCTTGGATACCAAGGA
Db	699	ACGCAACTTTACAGTGTCTATGGCAACTTTAACTGTATTAGCTCCTCTTCAITTTGTA	758	Qy	2094	GAAAGATGATGAATGACCACTGACCGTGTAGTCTTGGAGGCGGTACGAGTGAAGTCTGAGC
Qy	1014	GGCAGCGGCGCTCTGCGGAGCGCGGCGCAGCACCGCGCTTCTGCTGGCGCGGCTC	1073	Db	1782	CCTACTCTATCTCTCAGTGGCTTAGACCCAGAGAAATATCATGTGAGACTCTCTGGC
Db	759	ATGGCCAGAAAGTTTAAACAGGCTCGAGCTGCGCACTGCTGATTTGTGTGTCAGGAGA	818	Qy	2154	TTTCAAACAAACAGCAGGCGCTACGCTGTGTGTGGAAAGGCAAGACGAGAGAGCGCC
Qy	1074	CGGAGGACCGCGCGCTGCACTGCTGCTGACAGCGGATCCGTTTGGACCCCAATGG	1133	Db	1842	TTACAACAACATAGACGATGGCTATCAGGCAGATCAGATGTCTCAGCAC---TCCAGGATG
Db	819	AGGAATCCCTCTCCAGATGTCTATGTTGAAATGGAAGAGATACATTCGATGG	878	Qy	2214	CAGCCAGACCTGCTTATCAGAGGGGCGCACCGCTGCTCTTGGCCCATGTCCACGAGA
Qy	1134	CGCGCTCAAGGTGAGGCGGTGGCGGAGCTTGGTCACTCATCTCAGATCGGCTGCGAGA	1193	Db	1899	CGTGTCTGTGTGATCGATGCTCTCTCCACACCAACCCACCCTCTTATGCGAA
Db	879	TAGAATTAATGATACACAGT-----AAATTGGTAATTAACAGATTAATTCCTGAAGA	932	Qy	2274	GTCAAAACAGCTCCACATTTTGGCTTCGGTGGAAAGAGCCAGACTTTTACCACTGTCAA
Qy	1194	CGCTGGCTACTACAGTGTGATGAGAAACAGCGCGGGAATCCCTGTGCGGCTGCGCC	1253	Db	1959	GGCTAAACCTCATCTTCCATCTTCTGCACTGGAGAGGCTGTGCAITTCACCGCTGACA
Db	933	TGATGCTATTTATCAGTGTGATGGCTGAGAAATAGCAAGGATCTATTTTATCTAGAGCCAG	992	Qy	2334	GATTGTCAACTACATCTGTACGCTTCGGCCCTTGGGGGCTCAGGAATGCTTCCCTGTGTCAC
Qy	1254	CCTGGCGTGTGTGCGAGGCGCTGCCAGCGCCCGACTCGGCTCAGCAGCCAGCC	1313	Db	2019	AATCAATTAACATACCACTCGCTGTAACTCTGTGTGGCTTGCAGAAATGCTTCTTGTGTCT
Db	993	ACTGACTGTGTGATGTGAGAAGACAGCCAGTGTCTCCCTATTAATGTACATGTGTAAC	1052	Qy	2394	CTACTATACAGCTCTGGAGAGACATCTCATTTGGCGGCTGAAACCAATTTACCAAGTA
Qy	1314	GCTGAGCAGCTCTCTGTGTGTGGCTGGGAGCGGCTGAGTTGCAACGAGCAAAAT	1373	Db	2079	GTACCTTCAAAACATCAGAAATCTCATATGTTGGTTCAAGGCTTAGAAACCAAAAT
Db	1053	CATGTCAAGCTCAGCCATCTTTTAGCTGGGAGAGGCCACTTTTATAATTCAGACAAAGT	1112	Qy	2454	CGAGTTTGGGTACAGTCCACGAGTGGATATGGATGGGCTCTTGGCTCCGCTGTAGTA
Qy	1374	CATTGGCTTCTCTTCTACCAAAAGGAGGAGTGGACAAATGTGGAGTACAGTT	1433	Db	2139	CGAATTTGGCGTTCGATTAATATGGAATCAGTCTTCAGTCTCTGGAGCCCTGTGTGCTTA
Db	1113	CATTGGCTTATCTGTACACTACATGAAAGCAGAGGTTTAAATTAATGAAGATATCAAGT	1172	Qy	2514	ACGCTCCACCTGCTGACCGGCTTCAACACCTCTTCTGCACTGTGCGCTGAGCCGCCCT
Qy	1434	TGCAGTAAACATACACACAGAGCTGCAAGTTTGGGAGCTGGAACCCACACGAGTTA	1493	Db	2199	CCATTCTACTCTTCCAGAGCACGAGCCCAACCAAGTTGGATGAAAGTGAATTAAT
Db	1173	AGTCATCGAAATGACACACTCATTAATTAATGATGACTTAGAGCTGCGACCAATTA	1232	Qy	2574	GACACCATCCCGCTTGGGTGTTACATGTGTGTGCTCCCAACGAGCCCAATGGTGAGATTGT
Qy	1494	TGAGTTTCTAGTGTGGCTACTCCAGCTGGGCGCAGCGCAACCTCCAGCCAGCCCT	1553	Db	2259	AGAGGATGACACTGCGCTGGTTCTTGGAAACCCCTGTATGGCCCGAGAAACAGTTGTGAC
Db	1233	TACTTCTACATTTAGCATATATGCCAATGGGAGCCAGCCAGATGTCTGACCATGTGAC	1292	Qy	2634	GGAGTATCTAATTTCTTACAGCAACACACACCGCCCGAACAACAGTGGACATCTGCT
Qy	1554	GGTGATACACTGAGCAGATGCCAGCGCAGCAACCCAGCTTACTTGTCCAGCCCA	1613	Db	2319	CCGCTATATCTTATATGTCATCTAGGAAGGCTTGAATTCAGGAGAGTGGCAGGCTT
Db	1293	ACAGAACTACTTAGAGGATGTTCCTCTGAGACCTCTGAAATTTAGTTTGAACAATCGAAG	1352	Qy	2694	CACACAGAGGGAACATCTTCAAGTGCAGAGGTCTTCAAGTGTAGAGTGCACACTCGGTA
Qy	1614	CCCTCGACATCAGGTTGGCTGCTGCCCTGCCCTCCAGCTGAGCAATGACAGGT	1673	Db	2379	ACACCGTGAAGGGAATTAACCATGCTTGTCTAGAAACTTGGTAGCAGGAATTTGTA
Db	1353	TCCCACTGATATTTCTCATCTCTGGCTGCCAATCCAGCCCAATATCGGCGGGGCCAAGT	1412	Qy	2754	TTTCTTCAAGATGGAGCCCGCAGAGGTGGGCGCTTGGCCCTTTTCCGCTTGCAGGA
Qy	1674	GCTGAAGTACAAGATAGAGTACGGTTTGGGGAAGAGATCAGGTTTTTCTCCACCGAGGT	1733			

Db 2439 CATTGTCAAGATATCTGCATCCAAATGAGTGGGAGAGGACCCCTTTTCAAAATCTGTGGA 2498
Qy 2814 TGTGATTAATCTGCGAAGAGACATTTCTCAGACTC 2846
Db 2499 GCTGGCAGTACTTCCAAAGGAAACCTCTGAATC 2531

RESULT 2
US-09-877-730-1
; Sequence 1, Application US/09877730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding the
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3210
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-1

Query Match 5.1%; Score 316.6; DB 4; Length 3210;
Best Local Similarity 47.4%; Pred. No. 1.1e-71;
Matches 1210; Conservative 1; Mismatches 1270; Indels 72; Gaps 6;

Qy 300 AGACAGCATGATGAGGAAGCTTTAGGATCTGGAAGGTCACTGAGGCGAGCTATTCTG 359
Db 279 ATACATCAGTGGTGGAGGCGGAGGAGGAGCAGTCCGATGAAGGATTTATCAGT 338
Qy 360 TCTGGCCACAGCCGCTAGAGTGGTGGCCAGCCAGGTTGTGTGGTCAAGCTTGCAC 419
Db 339 CTGGCAATGAACAAATAGAGGCAATTTAGTCAAAAGACTCATCTTGCCTTATCAAC 398
Qy 420 ACTCGAAGACTTCTCTGACACCCGAGTCCAGATTTGTGAGGAGAACGGGACAGCAG 479
Db 399 TATTCTGCATTTGAGTTCAGCCCAATTTCCACTGAGTCCAGAGTGGAGTTGCTCG 458
Qy 480 CTTTGAATGCCACCAAGGCGCTTCCAGCCCGCATCATTTACTTTGGGAAAGGACCGGT 539
Db 459 ATTTGCATGCAAGATTTTCATCCACCCCTCTGCAGTCATACATGAGGATTTCAATCGAC 518
Qy 540 GACCGTGCCTGAG---GAGCCCGGCTCATCATCTTCCAGTGGCTTCCAGATCCT 596
Db 519 AACTCTACCTATGACTATGAGCAGGATACTGCCCCATCAACAGGAGTATTCAGATCTA 578
Qy 597 AGATGTCAGGACAGTGAAGTCTCTACCGTGGTGGCCACCAATTCAGCCCGCCA 656
Db 579 TGATGTCAGCCAAAGGATTTCTGMAAATATCTGTTATTTGCTGCCACTGTAGGCCACG 638
Qy 657 ACGATTACGAGGAGGCGCTCGCTCATCTGTGGCCCTCAGAGGGTCTTTGGAGGCTACCA 716
Db 639 ACGTAAAGTATGGAGGCGCTCGCTAACTGTGATTCAG---CTAAGGAGTCAAAATC 692
Qy 717 GGGGAGGATGTTGATTTGGCGCCCGCCAGAGAACACCAAGGTAGTGTCTGGACAGAA 776
Db 693 CTTCCACACACCAACAAATATAGCAGGTCCACAGAAACATAACAACATCTCTTTCACAG 752
Qy 777 TGATGATGAGTGGGCTCTGTGACCCCGCCCTTTTGTCTCGGTGGGTCCGACA 836
Db 753 TGTAGTTTGGAAATGATGGCCACAGGAAATCCCAAAACCAATCATTTCTTGGAGCCGCT 812

Qy 837 GGATGGAAAGCCTAT---CTCCACGGATGTCTGTTCTGGGCGGACCAATCTACTCAT 893
Db 813 TGATCACAAAATCCATTTGATGTCTTTAATACTCGGGTACTTTGGAAATGGTAATCTCATGAT 872
Qy 894 CGCAGGCGCGCAGCCTCGGCACCTCTGGAGTCTATGTCTGCGGAGCCCAACAGCCCTCCAC 953
Db 873 ATCTGATGTCAGGCTACAAATGCTGGAGTATATGTTTGTGCGGCCACTACCTCTGGCAC 932
Qy 954 GCGTGACTTCCCACTGCGGCTGCTGAGTCTCGAGTCTCGAGTCTTGTCTGCCCGCAGCATCTCGCA 1013
Db 933 ACGCAACTTTACAGTGTCTATGCGCACTTTAACTGTATTAGTCTCTCTTCAATTTGTTGA 992
Qy 1014 GGCACCGAGGCGCTCTCGGAGCGGGGCCAGCAGCCGCGCTTCTGTGCGGGGCTC 1073
Db 993 ATGGCCAGAAAGTTTAAACAAGGCGCTCGAGCTGGCACTGCTCGATTTGTGTGTACGGCAGA 1052
Qy 1074 CCGGAGGCCACGCGCGCTGCACTGCTGCGACAGCGGATCCGTTTGGACCCCAATGG 1133
Db 1053 AGGAATCCCTCTCTCCCAAGATGTCTATGTTGMAAATAGGAAGAGATACATTCGAATGG 1112
Qy 1134 GCGCGTCAAGGTGACGCGGCTGGCGCACTTGGTCACTCATCTCAGATCGGCTTGCAGGA 1193
Db 1113 TAGAAATTAATAATGTACAAACAGT-----AAATGGTAAATTAACCCAGATTTATCTGAAGA 1166
Qy 1194 CGCTGGCTACTACCAAGTGGTAGCAGAAACAGCGCGGGAACCTGCTGTGCGCGTGGGCC 1253
Db 1167 TGATGCTATTATTCAGTGCATGGCTGAGAATAGCAAGGATCTATTTATCTAGAGCCAG 1226
Qy 1254 CTTGGCGGTAGTGTGCGCGAGGCGCTGCCAGCGCCCGACTCGGGTCAAGCCAGCCGCC 1313
Db 1227 ACTGACTGTAGTGTGTCAGAAAGACAGACCCAGTGTCTCCCTATTAATGTATCATGTGAAAC 1286
Qy 1314 GCTGAGCAGCTCTCTGTGTGTGGCTGGGAGCGGCTGTGAGTTGTCACACGAGCAAAAT 1373
Db 1287 CATGTCAAGCTCAGCCATCTCTTTAGCTTGGAGAGGCCACTTTTATTAATTCAGACAAAGT 1346
Qy 1374 CATTTGGCTTCTCTTCACTACCAAAAGGACAGGGGAGTGACAAATGTGAGATACCAAGT 1433
Db 1347 CATTTGCCCTATTCTGTACACTACATGAAGCAGAGGTTTAAATTAATGAAGAGTATCAAGT 1406
Qy 1434 TGCAGTAAACAAATGACACACAGAGTGTGGGAGTGGGAGCTGGAAACCCCAACAGGATTA 1493
Db 1407 AGTCATCGGAAATGACACAACTCATTAATTAATGATGACTTAGAGCTTGCAGCAATTA 1466
Qy 1494 TGAGTTCTAGTGTGGCTTACTCCAGCTGGGGGCGAGCGAAGCTTCCAGCCCGAGCCCT 1553
Db 1467 TACTTTCTACATTTGTAGCATATATGCCAAATGGGAGCCAGCCAGATGTCTGACCATGTGAC 1526
Qy 1554 GGTGATACACTGGAAGATGTCCCGAGCGGACGACCCAGCTTACCTTGTTCAGCCCGCAA 1613
Db 1527 ACAGAAATCTCTAGAGGATGTTCCTCTGAGACCTCTCTGAAATTAAGTTTGAACAGTCCGAG 1586
Qy 1614 CCCTCGGACATCAGGGTGGCATGGCTGCGCCCTGCTCCAGCTTCCAGCTGAGCAATGAGCAGGT 1673
Db 1587 TCCACATGATTTCTCATCTCTGCTGCTGCAATCCCAAGCCAAATATCGGGGCGGCGCAAGT 1646
Qy 1674 GCTGAAGTCAAGATAGAGTACGTTTGGGGAAGGAGATCAGGTTTCTCCACCGAGGT 1733
Db 1647 GGTGCTGTATGCTTGTCTTTCGCTTAAGTACTGAGAAATTCATTCAGATTTCTGAGAGCT 1706
Qy 1734 GCCTGGAAATGAGACACAACTTACGTTTAAACTCACTTTCAGCCCAACAAAGTGTACCGAGT 1793
Db 1707 CCGGGGACCAACATGAGTACCTTTTGGAAAGCCCTGAAACCTGACAGTGTCTACCTGGT 1766
Qy 1794 CCGGATTTAGCTGGCATCTGGCTGCTGATGAGTCCCTTCTCAGTGGATGACAGCAG 1853
Db 1767 TCGGATTTACTGCTGCCACAGAGTGGGGCTGGGAGTGCATCATATGAGTATGAGTTCACATAG 1826
Qy 1854 GACACCTGGTGTGCACAAACAGAGCCATGTTTCCCTTTGCGCCCTGCAAAATTTGAGGTTGAG 1913
Db 1827 GAGCCCAAGAGCTACAAAGCGTGAAGCCCTTAAGTCTCCAGAGTTGCAITTTGGAGCTCT 1886
Qy 1914 GGCAAAAGATGGAGTCCCTGCTGTGTGTGATGCGAGCCGCCCTCCACCCACCCAGATCTC 1973

```
Db 1887 GAACGTACCAACCACTTTCTGTGAGGTGGCAGCAAGATGTAGAGGACACAGCTGCTATTCA 1946
Qy 1974 TGGATACAAACTCTACTTGGGAGAGTGGGAAACAGAGGAGGAGGAGAGTGGTACGCCCC 2033
Db 1947 GGGCTACAAAGCTGTACTACAGGAAGAAGGCGACGAGAGAA-----1988
Qy 2034 CCAGGGGGTCTGTGGAGATCAAGCTTTGGGACGTCTGGGCGCGCTGCGCTGGAAGAAGAAAGT 2093
Db 1989 -----TGGGCCCACTTTTCTTGGATACCAAGGA 2015
Qy 2094 GAAGCAGTATGAATCAGCACCACTAGTCCCTGGCAGCGCGCTAGAGGTGAAGTCTGTAGC 2153
Db 2016 CTTACTCTACTCTCTAGTGGCTTAGACCCCAAGAAAATATATCTGTGAGACTCTCTGGC 2075
Qy 2154 TTTCAACAACACAGAGAGGCTACGCTGTGTGTGAGAGGCAAGCGGAGAGGCGCC 2213
Db 2076 TTACAACAACATAGACATGGCTATCAGGCGAGATCAGACTGTGAGCAC---TCAGGATG 2132
Qy 2214 CAGCCAGACCTCGCTATCAGAGGGGGCCACCGCTCGCTCTGCCCCTATGTCCACGCAGA 2273
Db 2133 CGTGTCTGTTCGTGATGCGATGCTCTCTCCACCAACCCACCACTCTCTATGGAA 2192
Qy 2274 GTCAACAGCTCCACTCTCAATTTGGCTTGGTGGAAAGAGCCAGACTTTTACCACTGTCAA 2333
Db 2193 GGCTAACACCTCATCTTCTCATCTTCTGCACTGGAGAGGCTGCAATTCACCGCTGCACA 2252
Qy 2334 GATTGTCACTACACTGTACGCTTGGCCCTGGGGCTCAGAAATGCTTCCCTGTCTAC 2393
Db 2253 AATCATTAATACACCATCGCTGTAACTCTGTGGCCCTGCAGAAATGCTTCTTGTGTTCT 2312
Qy 2394 CTACTATACAGCTCTGGAGAGACATTTCTCATTTGGCGGCTGAAACCACTTTTACCAAGTA 2453
Db 2313 GTACCTTCAACATCAGAACTCAGATTTGGTTCAAGTCTTAGAACCAACCACTAATA 2372
Qy 2454 CGAGTTTGGGGTACAGTCCCAAGAGTGGATATGAGTGGGCCCTTTTGGCTCGTGTAGA 2513
Db 2373 CGAATTTGGCGTTTGGATATCATGTGATCAGCTTTCCAGTCTTGGAGCCCTGTAGTCTA 2432
Qy 2514 ACCTCCACCTCGCTGACCGGCTTTCAACCTCTCTGACCTGGGCTGAGCCCT 2573
Db 2433 CCATTTCTACTCTTCCAGAGACACAGCAGGCCCCACCAAGTTGGAGTAAAGTGAATTAAT 2492
Qy 2574 GACACCATCCACGTTTCGTTTACATGTTGTCTGCTGCTCCCGAGGCCCAATGCTGAGATTGT 2633
Db 2493 AGAGATGACACTGCTGCTGTTTCTTGGAAACCCCTGTAGTGGCCAGAAACAGTTGTGAC 2552
Qy 2634 GAGATATCTAATTTCTTACAGCAACCAACCAACCCAGCCCGAACAACAGTGGACACTGGTA 2753
Db 2553 CCGCTATACTATCTTATATGCACTAGGAAGGCTGGATTTGCAGGAGAGTGGCAGGTCTT 2612
Qy 2694 CACCAAGAGGGAACATCTTCAAGTGCAGAGTCCATGGCTAGAGAGTGAACACTGGTA 2753
Db 2613 ACACCGTGAAGGGGCAATAACCAATGGCTTTGCTAGAAACTTGGTAGCAGGAAATGTGA 2672
Qy 2754 TTTCTTCAAGATGGAGCCGCGCAGAGGTGGGCGCTGCGCCCTTTTCCGCTTTCAGGA 2813
Db 2673 CATTTGACAGATATCTGATCTCAATGAGTGGGAGAGGACCCCTTTTCAAAATTTCTGTGA 2732
Qy 2814 TGTGATTACTCTGCAAGAGACATTTCTCAGACTC 2846
Db 2733 GCTGGCAGTACTTCCAAAGGAACCTCTGAATC 2765
```

RESULT 3

```
US-09-877-730-17
; Sequence 17, Application US/0987730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
```

```
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632e1 Human Phosphatases and Polynucleotides Encoding
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3219
; TYPE: DNA
; ORGANISM: homo sapiens
; US-09-877-730-17
```

```
Query Match 5.1%; Score 316.6; DB 4; Length 3219;
Best Local Similarity 47.4%; Pred. No. 1.1e-71;
Matches 1210; Conservative 1; Mismatches 1270; Indels 72; Gaps 6;
```

```
Qy 300 ACACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTCG 359
Db 45 ATACATCAGTGAAGTGAAGGCGGAGGAGAGCAGTCCGATGAAGGATTTTATCAGTG 104
Qy 360 TCTGGCCACACCCCGCTAGGAGTGGTGGCCAGCCAGGTTGTGTGGTCAAGCTTGGCCAC 419
Db 105 CTTGGCAATGAACAAATATGAGGCCATTTAGTCAAAAAGCTCATCTTGCCTTATCAAC 164
Qy 420 ACTCGAAGACTTCTCTCTGACCCCGAGTCCAGATTTGTGGAGGAGAACGGGACAGCAG 479
Db 165 TATTTCTGCATTTGAAGTCCAGCCAAATTTCCACTGAGGTCCACGAAGTGGAGTTGCTCG 224
Qy 480 CTTTGAATGCCACACAGGGGCTTCCAGCCCGCTCATCTTCCAAAGTGGCTCTCCAGATCCT 539
Db 225 ATTTGATGCAAGATTTTCATCCACCTCTCGAGTCATAACATGGAGTTCATATCGGAC 284
Qy 540 GACCGTCCCTGAG---GAGCCCGGCTCATCTCTTCCAAAGTGGCTCTCCAGATCCT 596
Db 285 AACTCTACTATGACTATGACAGGATACTGCTTACCACAGGATATTGAGATCTTA 344
Qy 597 AGATGTCCAGGACAGTGAATGAGGTCTTACCGCTGCGTGGCCACCAATTCAGCCGCCA 656
Db 345 TGATGTCCAGCAAAAGGATTTCTGGAATTTATCGTTGTTATTTGCTGCACTGTAGCCACCG 404
Qy 657 AGATTTACGACAGAGGCTCGCTCACTGTGGCCCTCAGAGGCTCTTGGAGGCTACCA 716
Db 405 ACGTAAAGTATGGAGGCTCTCGTAACTGTGATTCAG-----CTAAGGAGTCAAAATC 458
Qy 717 GGGGCGAGGATGTGGTCAATTTGGCAGCCCGCAGAGAACACACGCTAGTGTCTGGACAGAA 776
Db 459 CTTCCACACACCAACATTTATAGCAGGTCCACAGAACATAACAACATCTTTCATCAGAC 518
Qy 777 TGTAGTGAATGAGTGGTGGCTCTGTGACCCCGCCCTTTTGTGTCTGTGGTCCGACA 836
Db 519 TGTAGTTTGGAAATGATGCGCCACAGGAATCCCAAAACCAATCTTCTTGGAGCCGCT 578
Qy 837 GGATGGAAGCCCTAT---CTCCACGAGTGTCTGTTCTTGGGCCGACCAATCTACTCAT 893
Db 579 TGATCAAAATCCATTTGATGTCTTTAATACTCTGGGTACTTGGAAATGGTAAATCTCATGAT 638
Qy 894 CGCAGCGCGCAGCGCTCGGCACCTCTGAGTCTATGTCTCGCAGGCAACAAGCCCTCAC 953
Db 639 ATCTGATGTCCAGCTACAAACATGTGGAGTATATGTTGTGGGCCCATCTACCCCTGGCAC 698
Qy 954 GCGTGAATTCGCCACTGCGGCTGTGAGCTCCGAGTGTGTGTGGCCCGACCACTCTCGCA 1013
Db 699 ACGCAACTTTACAGTTGCTATGCGAACTTTTAACTGTATTAGCTCTCTCTTCAATTTGTTGA 758
Qy 1014 GGCACCCGAGGCGCTCTCGCGACGCGGCCGACGACCGGCGCTCTGTGTGCCGGCGTC 1073
Db 759 ATGGCCAGAAAGTTTAAACAAGGCTCGAGCTGCGACTGCTCGATTTGTGTGTCAGGCAGA 818
```

QY 1074 CGGGAGCCACGGCCCGCTGCACTGCTGCAAGACGGGATCCGGTGGACCCAAATGG 1133
DB 819 AGGAATCCCTCTCCCAAGATGTCAATGGTTGAAAATGGAAGGAGATACATTCGAAATGG 878
QY 1134 CGCGTCAAGGTGACAGGCGGTGGCGGAGCTTGGTCACTCACTCAGATCGGCTGCAGGA 1193
DB 879 TAGAATTAATATGACACAGT-----AAATGGTAATTAACAGATATTCTCTGAAGA 932
QY 1194 CGTGGCTPACCAAGTGGTAGCAGAAACAGCGGGGAACTGCTGTGCGCTGCGCC 1253
DB 933 TGATGCTATTATCAGTGCATGGCTGAGATAGCCAGGATCTATTTTATCTAGAGCCAG 992
QY 1254 CCTGGCGTAGTGTGGCGAGGGGCTGCCAGGCGCCGAGCTGGGTACAGCCAGCC 1313
DB 993 ACTGACTGTAGTGTGACAGAGACAGACCCAGTGTCTCTATTAATGTACATGCTGAAAC 1052
QY 1314 GCTGAGCAGCTCTCTGTGCTGTGGCTGGAGCGGCTGAGTTGGACAGCGAGCAAT 1373
DB 1053 CATGTCAAGCTCAGCAATCTTTTAGCTGGAGAGGCCACTTTATTAATTCAGACAAGT 1112
QY 1374 CATGGCTTCTCTCTTCACTACCAAGGCAAGGGAGTGACAAATGTGGAGTACCAATT 1433
DB 1113 CATGGCTATTCTGTACACTACATGAAAGCAGAGGTTTAAATTAATGAAGATATCAAGT 1172
QY 1434 TGCAGTAAACAAATGACACACAGAGCTGCAGGTTGGGACCTGGAAACCAACAGGATTA 1493
DB 1173 AGTCATCGAAATGACACAACTCATATTAATTAATGATGACTTAGAGCTGCACGAATTA 1232
QY 1494 TGAGTTCTACGTGTGGCTACTCCAGCTGGGGGCGAGCGAACCTCCAGCCCGCCCT 1553
DB 1233 TACTTTCTACATGTAGCATATATGCAATGGAGCGCAGCAGATGTCTGACCATGTGAC 1292
QY 1554 GGTGATACACTGGACGATGCTCCAGCGCAGCAGCCCGAGCTTACCTTGTCCAGCCCAA 1613
DB 1293 ACAGAACTACTAGAGGATGTTCCCTGAGACCTCTCTGAAATTTAGTTGACAACTGAG 1352
QY 1614 CCCCTGGACATCAGGGTGGCATGCTGCTCCCTGCTCCAGCTGAGCAATGACAGGT 1673
DB 1353 TCCCACTGATATTCTCATCTCTGCTGCAATCCAGCCAAATATCGCGGGGCCAAGT 1412
QY 1674 GCTGAAGTACAGATAGATGATGTTGGGAGAGGAGATCAGGTTTCTCCAGCGAGT 1733
DB 1413 GGTGCTGTATCGCTGTGCTTTCCGCTTAAGTACTGAGAAATTCATCCAAAGTTCTGGAGCT 1472
QY 1734 CCCTGGAAATGAGACACAACTTACGTTAACTCACTTCAGGCCAAACAAAGTGATCCAGT 1793
DB 1473 CCCGGGACACGATGATGATCTTTGGAGGCTGAAACCTGACAGTGTCTACCTGGT 1532
QY 1794 CCGAATTCAGCTGGCACTGGCGTGGCTATGGAGTCCCTTCTCAGTGGATGACAGCAG 1853
DB 1533 TCGAATTAAGTCCACAGAGTGGGCTGGAGAGTCACTAGTATGGACTTCACATAG 1592
QY 1854 GACACCTGGTGTGACACACAGAGCATGTTCCCTTTGCCCCCTGCAGNAATGAAGGTGAG 1913
DB 1593 GACGCCCAAGACTCAAGCTGGAAGGCCCTTAAGTCTCCAGAGTTGCAATTTTGGAGCTCT 1652
QY 1914 GGCAAGATGGAGTCCCTGTGTGTGTCATGTCAGCGCGCCCTCCACCCACCCAGATCTC 1973
DB 1653 GAACGTGACCACTTTCTGTGAGGTGGCAGACAGATGTAGAGCAGACAGCTGCTATTCA 1712
QY 1974 TGGATACAACTCTACTGGGAGAGGTGGGAACAGAGGAGGAGGAGGAGTGTGACCGCCC 2033
DB 1713 GGGCTACAAGCTGTACTACAGGAAGAGGGCAGCAGAGAA-----1754
QY 2034 CCCAGGGGTGCTGGAGATCAAGCTTGGGACGTGGGCGCCCTGGCGCTGAAGAGAAAGT 2093
DB 1755 -----TGGGCCCATTTTCTTGGATACCAAGGA 1781
QY 2094 GAACAGTATGAATGACCCAGTGTAGTCCCTGGCAGGCGGTACGAGGTGAAGCTCGTAGC 2153
DB 1782 CCTACTCTATCTCTCAGTGGCTTAGACCCCGAGAGAAATATCATGTGAGACTCTGGC 1841
QY 2154 TTTCAAAACACGAGGACGGCTAGCTGTGTGTGGAGGGGCAAGACGGAGAGGGGCC 2213

DB 1842 TTCAACAAACATAGACGATGGCTATCAGGCAGATCAGACTGTTCAGCAC-----TCCAGGATG 1898
QY 2214 CAGCCAGAGCTGCTTATCCAGAGGGGGCAGCGCTGCTCTCCCTCCATATGTCACGAGGA 2273
DB 1899 CGTGTCTGTGTGATCGCATGGTCCCTCTCCACACACACCCACCACTCTCTATGGAA 1958
QY 2274 GTCAACAGCTCCCACTTCCATTTGGCTTCGGTGGAGAGGAGCAGACTTTTACCACTGTCAA 2333
DB 1959 GGGTAAACACCTCATCTTCCATCTCTCTGCACTGGAGAGGGCCTGCATTCACCGCTGACA 2018
QY 2334 GATTTGCAACTACACTGTAGCTTGGCCCTGGGGCTCAGGATGCTTCCCTGGTTCAC 2393
DB 2019 AATCATTAATACACCATCCGCTGTATCTGTTGGCTCGCAGAAATCTTCTTTGGTCT 2078
QY 2394 CTACTATACAGCTCTGGAGAGACATCTTCATTTGGGGGCTGAAACATTTTACCAAGTA 2453
DB 2079 GTACCTTCAACATCAGAACTCAGATGTTGGTTCAGGTCTAGAACCAACACCAATA 2138
QY 2454 CGAGTTTGGGTAAGTCCCAAGGAGTGGATATGGATGGGCCCTTTGGCTCCGTCTGAGA 2513
DB 2139 CGAATTTGGCTTCGATTACATGTGGATCAGCTTTCCAGTCTTGGAGCCCTGTAGTCTA 2198
QY 2514 ACGTCCACCTGCTGACCGGCTTCAACACCTCTCTGACCTGGCGCTGAGCCCCCT 2573
DB 2199 CCATTTCTACTCTTCCAGAGACACAGCAGGCCCAACAGTTTGGAGTAAAGTGACATTAAT 2258
QY 2574 GACACCATCACCGTTCCGTTTACACTGCTGCTGCCCGCCACCGAGCCCAATGTTGAGATTGT 2633
DB 2259 AGAGATGACACTGCTGCTGTTTCTTGGAAACCCCTGTATGGCCAGAAACAGTTGTGAC 2318
QY 2634 GAGATATCTAATTTCTTACAGCAACACACACCCAGCCCGAAACACAGTGGACACTGCT 2693
DB 2319 CCGCTATACTACTTATATGATCATCTAGGAGGCTGTGATTCAGAGAGTGGCAGGTCTT 2378
QY 2694 CACACAGAGGGAACATCTTTCAGTGCAGAGGTTCATGGCTAGAGAGTGCACACTCGGTA 2753
DB 2379 ACACCGTGAAGGGCAATAACCATGCTTTGCTAGAAAACTTGTAGCAGGAATGTGA 2438
QY 2754 TTTCTCAAGATGGGAGCCCGCAGAGGTGGGGCTTGGGCCCTTTTCCCGCTTGCAGGA 2813
DB 2439 CATTTCAAGATATCTGCATCCCAATGAGGTGGGAGAGGACCCCTTTTCAAATCTGTGGA 2498
QY 2814 TGTGATTAATCTGCAAGAGACATTTCTCAGACTC 2846
DB 2499 GCTGGCAGTACTTCCAAAGGAACCTCTGAATC 2531

RESULT 4

US-09-877-730-31
; Sequence 31, Application US/09877730
; Patent No. 6455632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 645632el Human Phosphatases and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 3874
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-31

Query Match	5.1%;	Score 316.6;	DB 4;	Length 3874;
Best Local Similarity	47.4%;	Pred. No. 1.2e-71;		
Matches 1210; Conservative		1; Mismatches 1270;	Indels 72;	

Qy	300	AGA	CAGCGATGATGAGGAAGCTCTT	AGGATCTT	AGGATCTT	GGAAAGGTCACTGAGGGCAGCTATT	CTCTG	359
Db	469	ATA	CATCAGTGGGTGGAAGCGCAGGCAGGAGAGCAGTCCCGATGAAGGATTTT	TATCAGTG	528			
Qy	360	TCT	GGCCCCACAGCCCGCTAGAGTGGGTGGCCAGCCAGGTTGCTGTGCTCAAGCTTGCAC	419				
Db	529	CTT	GGCAATGAACAATAATGGAGGCATCTT	TAGTCAAAAAGCTCATCTTGCCCTTATCAAC	588			
Qy	420	ACT	CGAAGACTTCTCTCTGCAACCCCGAGTCCCAAGATTTGTGGAGGAGAAACGGGACAGCACG	479				
Db	589	TA	TTTCTGCAATTTGAAGTCCAGCCAAATTTCCACTGAGTCCACGAAGGTGGAGTTGCTCG	648				
Qy	480	CTT	TGAATGCCACCAAGGCGCTTCAGGCCCCCATCATTTACTTTGGGAAAGGACAGGT	539				
Db	649	ATT	TGCATGCAAGATTTTCATCCCAACCTCTCGCAGTCAATCAACATGGGAGTTCAATCGGAC	708				
Qy	540	GAC	CGTGCCTGAG--GAGCCCGGCTCATCACTTCTCCCAAGTGGCTCTCCAGATCCCT	596				
Db	709	AAC	TCTACCTATGACTATGACACAGGATACTGCCCTACCAACAGGAGTATTTGCGATCTA	768				
Qy	597	AGAT	GTCCAGGACAGTATCAGGCTCTTACCGTGGTGCGGCCACCACTAAATTCAGCCCGCCA	656				
Db	769	TGAT	GTGAGCCAAAGGATTTCTGGAATTAATCGTTGTTATTTGCTGSCCACTGTAGCCACCG	828				
Qy	657	ACG	ATTGAGCCAGGAGCGCTCGCTCACTGTGGCCCTCAGAGGTCTTTGGAGGTACCA	716				
Db	829	ACG	TAAAGATATGGAGGCGCTCGCTAACTGTGATTTCCAG-----CTAAGGAGTCAAAATC	882				
Qy	717	GGG	CGAGATGTGCTCATCTGTGGCAGCCCGACAGAAACACACGGTAGTGTCTGACAGAA	776				
Db	883	CTT	CCACACACCAACCAATTAAGCAGTCCACAGAAATAAACAATCTCTTATCAGAC	942				
Qy	777	TGT	AGTATGAGATGCGTGGCGCTCTGTGACCCACCCCTTTTGTGTCTTGGGTTCGACA	836				
Db	943	TGT	AGTTTGGAAATGCAATGGGCCACAGGAAATCCCAAAACCAATCATTTCTTTGGAGCGCCT	1002				
Qy	837	GGAT	TGGAAGCGCTAT---CTCCAGGATGTCACTGTTCTCGGGCGGACCAATCTACTCAT	893				
Db	1003	TGAT	CACAAAATCCAAATGATGTCTTTAATACTCGGGTACTTTGGAAATGGTAATCTCATGAT	1062				
Qy	894	CGC	CAGCGCAGCGCTCTCGGCACCTCTGGAGTCTATGTCTGCGGAGCCAAACAGCCCTCAC	953				
Db	1063	ATC	TGATGTGAGGCTACAACAATGTGGAGTATATGTTTGTGCGGCCACTACCCCTGGCAC	1122				
Qy	954	GCG	TGACTTCGCCACTGCGGCTGTGAGTCCCGAGTCTTGTGTGCCCCAGCATCTCGCA	1013				
Db	1123	ACG	CAACTTTTACAGTTGCTATGGCACTTTAACTGTATTAGTCTCTCTCTTCAATTTGTGA	1182				
Qy	1014	GGC	ACCGGAGCGCTCTCGCGAGCGCGGCGCAGCACCGCGCTTCGTGTGCGGGGCTC	1073				
Db	1183	ATG	CCAGAAAGTTTAACAAGGCTCGAGCTGGCATCTGCTGATTTGTGTGTCAGGCAGA	1242				
Qy	1074	CGG	GGAGCCACGCGCCGCTGTGCACTGGCTGCAAGCGGATCCGTTGCGAGCCCAATGG	1133				
Db	1243	AGG	AATCCCCCTCTCCCAAGATGTCAATGTTGAAAAATGGAAGGAGATACATTCGAATGG	1302				
Qy	1134	GCG	GTCAAGGTGCAAGGCGGTGGCGGACGCTTGGTCACTCACTCAGATCGCCCTGCAGGA	1193				
Db	1303	TAGA	TTTAAATGTACAACAGT-----AAATTGGTTAATTAACAGATTAATTTCTGAAGA	1356				
Qy	1194	CGC	TGGCTACTACCAAGTCCGTAGCAGAAAAACAGCGCGGAACTGCCTGTGCGCTGCGCC	1253				
Db	1357	TGAT	GTCTTATTAACAGTGCATGCTGAGAAATGCCAAGGATCTATTTTATCTAGAGCCAG	1416				
Qy	1254	CTT	GCGGTATGTGTGCGAGGCGGTGCCCCAGCGCCCGGACTCGGGTCAACAGCCAGCC	1313				
Db	1417	ACT	GACTGTAGTGTATGTACAGAAGACAGACCCAGTGTCTCCCTAATATCTATCTGTAAC	1476				

Qy	1314	GCTGAGCAGAGCTCCTCTGTGTCTGTGTGGGAGCGGCTGTGATTCGCACACGCGAGCAAAAT	13173
Db	1477	CATGTCGAAGCTCAGCCCAATCTTTTAGCTGTGGGAGAGGCCCCCATTTTATAATTTCAGACAAAGT	15316
Qy	1374	CATTGGCTTCTCTCTTCACTTACCACAAAGGCAAGGGGAGTGGACAATGTGAGGTACCAAGTT	14333
Db	1537	CATTGGCTTCTCTGTACACTACATGAAGCGAGAAAGTTTAAATATATGAAGAGTATCAAGT	15916
Qy	1434	TGCAGTAAACAATGACACCAACAGAGCTGCAAGTTTCGGGACCTCGGAAACCCCAACACGGAATTA	14913
Db	1597	AGTCATCGGAAAATGACACAACTCAATTATATTAATTGATGACTTATAGAGCTGCCAGCAATTA	16516
Qy	1494	TGAGTTCTACGTGGTGGCCTACTCCAGCTGGGGGGCAGCCGAACTTCAGGCCAGCCCT	15533
Db	1657	TACTTTCTCAATTTGTAGCATATATGCCAATGGGAGCCAGCCAGATGTCTTGACCATGTGAC	17116
Qy	1554	GGTGCATACACTGGACGATGTCCCAAGCGCAGACACCCAGCTTACCTTGTTCAGGCCCAA	16113
Db	1717	ACAGATATCTTAGAGGATGTTCCCTGAGACTCTCTGAATTTAGTTTGACATGTGGAAG	17716
Qy	1614	CCCTCGGACATCAGGGTGGCATGGCTGCCCTGCCCTCCAGCTCCAGCCCTGAGCAATGGAAGGT	16713
Db	1777	TCCCACTGATATTTCTCATCTCTCTGGCTGCCAAATCCAGCCAAATATATCGCGGGCCCAAGT	18316
Qy	1674	GCTGAAGTACAAGATAGAGTACGGTTTGGGGAGGAAGATCAGGTTTTTCTCCACCAAGGT	17313
Db	1837	GGTGTGTATCGCTTGTCTTTTCGCCCTAAGTACTGGAATTTCAATCCAAGTTCTGAGGCT	18916
Qy	1734	GCCTGGAAATGAGACACAACTTACGTTTAACTCACTTCAGCCAAACAAAGTGTACCGAGT	17913
Db	1897	CCCGGGACCAACGATGAGTACCTTTTGGAAAGCCCTGAACCTTGACAGTGTCTACCTGGT	19516
Qy	1794	CCGGAATTCAGCTGGCACTGGCGCTGGCTATATGAGTCCCTTCTCAGTGGATGACGACACAG	18513
Db	1957	TCGGATTACTGCTGCCACAGAGTGGGGCTGGGAGAGTCATCATGATGAGCTTCACATAG	20116
Qy	1854	GACACCTGGTGTGCACAAACAGAGCCCATGTTCTCTTGGCCCCCTGCAGAAATTGAAGGTGAG	19113
Db	2017	GAGCCCAAGCTACAAAGCGTGAAAGCCCCCTAAGTCTCCAGAGTTGCAATTTGGAGCCTCT	20716
Qy	1914	GGCAAGATGGAGTCCCTGGTGTGTTCATGGCAGCGCCCTCCACCCCAACCAAGATCTC	19713
Db	2077	GAACTGTACCACTTTCTGTGAGGTGTGACGACAGATGTGAGGACACAGCTGCTATPTCA	21316
Qy	1974	TGGATACAAACTCTACTGGGAGAGGTGGGAAACAGAGGAGGAGGCGCAGATGGTGACGCCCC	20313
Db	2137	GGGCTACAAGCTGTACTACAAGGAGAAAGGGCAGCAGGAGAA-----2178	
Qy	2034	CCCAAGGGGTCTGGAGATCAAGCTTGGGACGTCTGGGCCCTGTGGGCTGGAAGAGAAAGT	20913
Db	2179	-----TGGGCCCATTTTCTTGATACCAAGGA2205	
Qy	2094	GAAGCAGTATGAATGACCCAGTTAGTTCCTGGCAGCGCTACGAGGTGAAGTCTGTAGC	21513
Db	2206	CCCTACTATACTCTCTAGTGGCTTTAGACCCCAAGAAATAATCATGTGAGACTCTCTGGC	22616
Qy	2154	TTTCAACAAACAGAGGACGGCTACGCTGTGTGTGAAGGGCAAGACGAGGAAGCGCC	22113
Db	2266	TTACAAACAATAGACGATGGCTATCAGGCAGATCAGACTGTCAAGC-----TCCAGATG	23216
Qy	2214	CAGCCAGACCTGCCTATCCAGAGGGGGCAACCGCTGCTCTCTGCCCATGTCCACGAGA	22713
Db	2323	CGTGTCTGTTCTGTGATCGCATGTGTCCCTCTCCACCAACCCACCATCTCTATGCGAA	23816
Qy	2274	GTCAACAGCTCCACTTTCATTTGGCTTTCGGTGGAGAGACGACGATTTTACCACTGTCAA	23313
Db	2383	GGCTAAACACCTCATCTTCTTCATCTTCTGTGCACTGAGGAGGCGCTGCATTCACCGCTGACA	24416
Qy	2334	GATTGTCAACTACACTGTACGCTTTCGGCCCTCGGGGCTCAGGAAATGCTTTCCCTGGTTCAC	23913
Db	2443	AATCATTAACATACACCACTCCGCTGTAATCTCTGTTGGCTTCAGAAATGCTTCTTTTGTCT	25016
Qy	2394	CTACTATACAGGCTCTGGAGAGACATTTCTCATTTGGGGGCTCGAAACCAATTTTACCAGTA	24513


```
QY 1554 GGTGCATACACTGGAGATGTCCCGAGCGCAGACCCCGAGCTTACCTGTCTCAGCCCAA 1613
D 1527 ACAGATACTCTAGAGGATGTTCCCTGAGACCTCTCTGAAATTTAGTTTGACAAGTGAAG 1586
QY 1614 CCGCTCGGACATCAGGGTGGATGCTGCTGCCCTGCTCCAGCTGAGCAATGGAGAGT 1673
D 1587 TCCCATGATATTTCTCATCTCTGGCTGCAATCCAGCAAAATATCGGCGGGGCAAGT 1646
QY 1674 GCTGAAGTACAAGATAGATGACGTTTGGGGAAGGAAGATCAGGTTTTCTCCACCGAGT 1733
D 1647 GGTGCTGATCGTGTCTTTCCGCTTAAGTACTGAGAAATTCATCCAAAGTTCTGAGCT 1706
QY 1734 GCGTGAATAGAGACAATCTTACGTTAACTCACTTCAGCCAAACAAAGTGTACCGAGT 1793
D 1707 CCGGGGACCAACGATGATGACCTTTTGAAGGCGCTGAAACCTGACAGTGTCTACTGCT 1766
QY 1794 CCGGATTCAGCTGGGACTGGCGCTAGTGGAGTCCCTTCTCAGTGTGATGAGCAGACAG 1853
D 1767 TCGGATTAAGTCTGCTGCCACAGAGTGGGCTGGGAGAGTCAATGATGAGACTTCAATAG 1826
QY 1854 GACACCTGGTGTGCAACACAGAGCCATGTTCCCTTTGCCCCCTGCAGAAATTCAGGTGAG 1913
D 1827 GACGCCAAAGCTACAAGGTGAAGCCCTTAAGTCTCAGAGTTGCAATTTGGAGCTCT 1886
QY 1914 GGCMAAGATGAGTCCCTCTGCTGTGTATGGCAGCGCCCTCCACCCACCCAGATCTC 1973
D 1887 GAAGTGTACCACTTCTGTGAGTGGCAGCAAGATGTAGAGGACACAGCTGCTATTCA 1946
QY 1974 TGGATCAAACTCTACTGGGAGAGTGGAAACAGAGGAGGAGGAGATGTTGACCGGCC 2033
D 1947 GGGCTACAAGCTGTACTACAAGGAAGAGGCGAGCAGAGAA----- 1988
QY 2034 CCAGGGGTCTGGAGATCAAGCTTGGGACGTGGGCGCGTGGGCTGAAGAAAGT 2093
D 1989 -----TGGGCGCAATTTCTTGGATACCAAGA 2015
QY 2094 GAAGCAGTATGAATCAACCTAGTTCCTGCGAGCGCTACGAGTGAAGTCTGTAGC 2153
D 2016 CTTACTCTATCTCTCAGTGGCTTAGACCCAGAGAAATATCATGTGAGACTCTCTGC 2075
QY 2154 TTTCAACAAACAGGAGGAGTACCTGCTGTGTGTGGAAGGGGCAAGCAGGAGGCGCC 2213
D 2076 TTACAACAAACATAGACAGTGGCTATCAGGAGATCAGACTGTGAGCAGC-----TCCAGGATG 2132
QY 2214 CAGCCAGACCTGCTATCCAGAGGGGCGACCGCTGCTCTGCGCAATGTCCAGCAG 2273
D 2133 CGTGTCTGTGTGATGCAATGCTCTCTCCACCAACCCACCCACCTCTATGCGAA 2192
QY 2274 GTCAACAGCTCCACTTTCATTTGGCTTGGTGGAGAGAGCCAGACTTTACCACTGTCAA 2333
D 2193 GGGTAACACTCATCTTCATCTTCTGCACTGGAGAGGCTGCAATTCACCGCTGCACA 2252
QY 2334 GATGTCAACTACACTGTAGCTTTCGGCCCTTGGGGCTCAGGAATGCTTCCCTGCTAC 2393
D 2253 AATCAATTAACATACACCATCCGCTGTAATCTGTTGGCTGCAAGATGCTTCTTGGTCT 2312
QY 2394 CTACTATACCGCTCTGGAGAGACATTTCTATTGGCGGCTGAACCACTTTACCAAGTA 2453
D 2313 GTACCTTCAACATCAGAACTCACAATGTGGTTCAGAGTCTAGAACCAACCAAAATA 2372
QY 2454 CGAGTTTGGGTGATGATGCAAGAGTGGATGAGTGGGCTTGGGCTCGGTCTGATGA 2513
D 2373 CGAATTTGGCTGCAATATCATGTGATCAGCTTTCAGTCTTGGAGCCCTGTAGTCTA 2432
QY 2514 AGCTCCACCTGCTGACCGGCTTCAACACTCTCTTGAACCTGCGCTGAGCCCTCT 2573
D 2433 CCATTTCTCTTCCAGAGACACAGCAGGCCCAACAGTTGGAGTAAAGTGACATTAAT 2492
QY 2574 GACACCATCCACGTTGCTTACATGTTGTTCCCCCAAGGAGCCCAATGGTGAATGT 2633
D 2493 AGAGGATGACACTGCTGCTGTTTCTTGGAAACCCCTGATGGCCAGAAACAGTTGTGAC 2552
QY 2634 GGAGTATCAATTTCTCTCAGGACAAACCAACCCAGCCCGAACCAGCTGGACACTGCT 2693
```

```
D 2553 CCGCTATACTATCTTATATGATCATAGGAAGCCCTGATTCAGAGAGTGGCAGGCTCT 2612
QY 2694 CACACAGAGGAAACATCTTCACTGAGAGTCCATGAGGCTTCAAGTGTGACATCTGCTA 2753
D 2613 ACACCGTGAAGGGCAATTAACATGAGCTTCTAGTAAATACTTGTAGCAGGAATGTGTA 2672
QY 2754 TTTCTTCAAGATGGAGCCCGCACAGAGTGGGCGCTTGGGCGCTTTTCCCGCTTGCAGGA 2813
D 2673 CATTTCAAGATATCTGATCCATGAGTGGGAGAGGACCCCTTTTCAAAATTTCTGTGA 2732
QY 2814 TGTGATTAATCTCTCAAGAGACATTTCTCAGACTC 2846
D 2733 GCTGCGAGTACTTCCAAAGGAAACCTCTGAATC 2765
```

RESULT 6

```
US-09-877-730-15
; Sequence 15, Application US/09877730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 2481
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-15
```

```
Query Match 3.5%; Score 217; DB 4; Length 2481;
Best Local Similarity 50.3%; Pred. No. 5.8e-46;
Matches 647; Conservative 1; Mismatches 621; Indels 18; Gaps 4;
```

```
QY 300 AGACAGCATGATGAGGAAGCTCTTAGGATCTGGAAGTCTCACTGAGGCGAGCTATTCCTG 359
D 45 ATACATCAGTGAAGTGGAGGCGAGGCGAGGAGCAGTCCGATGAAGGATTTTATCAGTG 104
QY 360 TCTGCCCCACAGCCCGCTAGGAGTGGTGGCCGAGGCTTGTCTGTGGTCAAGCTTGGCCAC 419
D 105 CTGGCAATGAACAATAATGAGGCCATTTCTTAGTCAAAAAGCTCATCTTGGCTTATCAAC 164
QY 420 ACTCGAAGACTTCTCTGCAACCCCGAGTCCAGATTTGTGGAGGAGAAAGCGGACAGCAG 479
D 165 TATTTCTGCAATTTGAAGTCCAGCCAAATTTCCAGTGTCCAGAGTCCAGAGTGGTGGTCTG 224
QY 480 CTTTGAATGCCACCAAGGGGCTTTCAGCCCCCATTAATTTCTGGGAAAAAGGACAGGT 539
D 225 ATTTGATGCAAGATTTTCATCCACCTCTCTGCACTATAACATGGGAGTTCAATCGGAC 284
QY 540 GACCGTGGCTGAG---GAGCCCCGGCTCATCACTCTTCCCAAGTGGCTCTCCAGATCT 596
D 285 AACTCTACTATGACTATGAGCAGGATAACTGCCCTACCAACAGGAGTATTTGACATCTA 344
QY 597 AGATGTCCAGAGCAGTGTGATGAGGCTCTTACCGCTGCGTGGCGCACCAATTTAGCCCGCA 656
D 345 TGAATGAGCCAAAGGATTTCTGGAAATTTGTTGATTTGTTGTCACATGTAGCCACCG 404
QY 657 AGGATTCAGCCAGGAGGCTCTGCTCATCTGTGGCCCTCAGAGGCTCTTTGGAGGCTACAG 716
D 405 AGTAAAGTATGGAGGCTCTGCTAACTGTG-----ATTCCAGCTAAGGAGTCAAAATC 458
```

```
717 GGGGAGGATGCTGCTCATTTGGGAGCCCGGAGAAACACACGAGTGTCTGACAGAA 776
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
459 CTTCCACACACCAATATAGCAGGTCACAGACATACACATCTCTTCATCAGAC 518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
777 TGTAGTATGAGTGTGGCTCTGTGACCCACCCCTTTTGTCTGGTCCGAC 836
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
519 TGTAGTTTGAATGCTAGGCGCAGAGAAATCCCAACCAATCATTTCTTGAGCGGCT 578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
837 GGATGAAAGCTAT--CTCCAGGATGTCATCGTTCTGGCGCGGACCAATCTACTCAT 893
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
579 TGATCACAATCCATTGATGCTTTAATCTCGGGTACTTGGAAATGGTAATCTCATGAT 638
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
894 CGCCAGCGCGAGCTCGGCTCTGAGTCTATGTCGCGAGCCCAACAGCCCTCAC 953
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
639 ATCTGATGTCAGGCTACAACTCTGGAGTATATGTTGTGGGCCACTACCCCTGGCAC 698
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
954 CGGTGACTTCGCACTGGGCTGTGAGCTCGAGTCTTGTGTCGCCAGCATCTCGCA 1013
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
699 ACGCAACTTTACAGTTGTATGGCACTTAACTGTATTAGTCTCTCTTCAATTTGTA 758
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1014 GGCACCCGAGGCTCTCGCGAGCGGGCCAGCACCGCGCTTCTGCTGCGCGGCTC 1073
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
759 ATGGCCAGAAAGTTAAACAGGCTCGAGCTGGCACTGCTCGATTTGTGTGAGCAGAA 818
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1074 CGGGAGCGCACGGCCGCTGCACTGGCTGCAACGAGCGGATCCCGTTGCGACCCCAATGG 1133
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
819 AGGAATCCCTCTCCAGATGTCTATGTTGAAATGGAAGGAGATACATTCGATGG 878
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1134 CGCGTCAAGGTGCAAGGCGGTGGCGAGCTTGGTCACTCAGATCGGCTGCAAG 1193
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
879 TAGAATTAATATGTAACAAGT-----AAATGGTAAATTAACACAGATTTATCTGAAAGA 932
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1194 CGCTGCTACTACGAGTGGTAGCAAGAAACAGCGGGGAGTGGCTGTGCGCTGCGCC 1253
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
933 TGATGCTATTTATCAGTGCATGGCTGAGATAGCCAGGATCTATTTTATCTAGAGCCAG 992
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1254 CTTGCGGTAGTGGTGGCGAGGGGCTGCCAGCGCCCGGCTCGGGTCAACAGCCAGCC 1313
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
993 ACTGACTGTAGTATGTCAGAGACAGACCCAGTCTCCTATATGATGATGTCGTAAC 1052
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1314 GCTGAGAGCTCTCTGTGCTGGTGGCTCGGAGCGGCTGATTTGCAAGCGAGCAAT 1373
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1053 CATGTCAAGCTCAGCCATCTTTTAGCTGGGAGGCGCACTTTATATTCAGACAAAGT 1112
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1374 CATGGCTCTCTCTTCACTACCAAGCAAGGAGGAGTGAGCAATGTGGAGTACCATT 1433
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1113 CATTCCTATTTCTGTACACTACATGAAGACGAGAGGTTTAAATTAATGAAGAGTATCAAGT 1172
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1434 TGCAGTAAACATGACACACAGAGCTGCAGTTTCGGACCTTGGAAACCCACACGATTA 1493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1173 AGTCATCGAAATGACAACTCATATATATTATGATGATCTAGAGCTTGCAGCAATTA 1232
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1494 TGAGTTCTACGTGTGGCTACTCCAGCTGGGGCCAGCGGAACTCCAGCCAGCCCT 1553
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1233 TACTTTCTACATTTAGCATATATGCCAATGGAGCCAGCCAGATGTCGACCATGTGAC 1292
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1554 GGTGATACACTGGAGAGTGTCCCGAC 1580
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1293 ACAGAATATCTAGAGGATGACCCAG 1319
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 7

```
US-09-877-730-5
; Sequence 5, Application US/09877730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
```

```
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2715
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-5
```

Query Match 3.58; Score 217; DB 4; Length 2715;

Best Local Similarity 50.38; Pred. No. 6.1e-46;

Matches 647; Conservative 1; Mismatches 621; Indels 18; Gaps 4;

```
QY 300 AGCAGCGATGATGAGGAAGCTCTTAGGATCTTGAAGGTCACTGAGGCGAGCTATTCTCGT 359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 279 ATACATCAGTGAGGTGGAAGCGAGCGAGGAGAGCAGTCCGATGCAAGGATTTTATCAGTG 338
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 360 TCTGGCCACAGCCCGCTAGAGTGTGGCCAGCCAGGTTCTGTGTCAAGCTTGGCAC 419
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 339 CTTGGCAATGAACAAATATGAGGCCATTCTTAGTCAAAAAGCTCATCTTGCCTTATCAAC 398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 420 ACTCGAAGACTTCTCTCTGACCCCGAGTCCAGATTTGTGGAGAGAACGGGACAGCAG 479
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 399 TATTTCTGCAATTTGAAGTCCAGCCAAATTTCCACTGAGGTCCAGAAAGTGGAGTTGCTCG 458
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 480 CTTTGAATGCCACCAAGGCGCTTCCAGCCGCCCATCAATTACTTGGGAAAGGACCAGGT 539
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 459 ATTTGCATGCAAGATTTTCATCCACCTCTGTCAGTATACATATGGAGTTCAATCGAC 518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 540 GACCGTGTGAG---GAGCCCCGGCTCATCACTCTTCCCAAGTGGCTCTCCAGATCCT 596
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 519 AACTCTACCTATGACTATGAGCAGGATAAATGCGCTTACCAACAGAGATTTATGCAATCTA 578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 597 AGATGTCAGGACAGTGTGAGGCTCTTACCGCTGGTGGCCACCAATTCAGGCCGCCA 656
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 579 TGATGTGAGCCAAAGGAGTCTTGAAGATTTATGTTGTTATGCTGCCACTGTAGGCCACCG 638
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 657 ACGATTCAGCCAGGAGCGCTCGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTTACCAG 716
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 639 ACGTAAAGATGATGAGGCGCTCGCTAACTGTG-----ATTCAGCTAAGGAGTCAAAATC 692
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 717 GGGCAGGATGTGCTCATTTGTGGCAGCCCGCAGAGAACACACGCTAGTGTCTGGACAGAA 776
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 693 CTTCCACACACCAACCAATATAGCAGGTCCACAGAACATACAAACATCTCTTCATCAGAC 752
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 777 TGTAGTATGAGTGTGGCTCTGTGACCCCGCTTTTGTCTGCTGGTCCGAC 836
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 753 TGTAGTTTGAATGCTAGGCGCAGGAAATCCCAACCAATCATTTCTTGAGCGCGCT 812
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 837 GGATGAAAGCTAT---CTCCAGGATGTCATGTTCTGGCGCGGACCAATCTACTCAT 893
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 813 TGATCACAATCCATTGATGCTTTAATCTCGGGTACTTGGAAATGGTAATCTCATGAT 872
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 894 CGCCAGCGCGAGCTCGGCACTCTGGAGTCTATGTCGCGAGCCAAACAGAGCCCTCAC 953
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 873 ATCTGATGTCAGGCTACAAACATGCTGGAGTATATGTTGTGGGCCACTACCCCTGGCAC 932
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 954 CGGTGACTTCGCACTGGGCTGTGAGTCCGAGTGTCTGCGCCAGCCAGCATCTCGCA 1013
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 933 ACGCAACTTTACAGTTGCTATGGCAACTTTAACTGTATTAGTCTCTCTTCAATTTGTA 992
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1014 GGCACCCGAGGCGCTCTCGCGAGCGGGCCAGCACCGCGCTTCTGCTGCGCGGCTC 1073
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 993 ATGGCCAGAAAGTTTAAACAGGCTCGAGCTGGCACTGCTCGATTTGTGTGAGCAGAA 1052
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1074 CGGGAGCGCACGGCCGCTGCACTGGCTGCAACGAGCGGATCCCGTTGCGACCCCAATGG 1133
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 1053 AGGAATCCCTCTCTCCCAAGATGTCAATGTTTGAATAATGGAAGGAAGATACATTCGAATGG 1112
Qy 1134 GCGCGTCAAGGTGCGAGGCGGTGCGCGCAGCTTGTTCATCACTCAGATCGGCTCGCAGGA 1193
Db 1113 TAGAATTAATGATCAACAGT-----AAATTTGGTAAATTAACAGGATTAATTCCTGAAGA 1166
Qy 1194 CGCTGGCTACTACAGTGCCTAGCAGAAAACAGCGCGGAACTGCCCTGTGCCGCTGCGCC 1253
Db 1167 TGATGCTATTATCAGTGCATGGCTGAGATAGCAAGGATCTATTTTATCTAGAGCCAG 1226
Qy 1254 CTGGCGGTAGTGTGCGGAGGGGCTGCCAGGCGCCCGCACTCGGGTCAAGCCAGCCGCC 1313
Db 1227 ACTGACTGTATGTATGTGAGAAAGACAGACCCAGTGTCTCCCTATAATGTATACATGCTGAAC 1286
Qy 1314 GCTGAGCAGCTCTCTCTGTCTGTGGCTGCGGAGCGGCTGAGTTGCACAGCGAGCAAT 1373
Db 1287 CATGTCAAGCTCAGCCATCTTTTAGCTGGAGAGGCCACTTTATTAATTCAGACAAAGT 1346
Qy 1374 CATTTGGCTTCTCTCTCACTACAAAGGCAAGGGAGTGGCAATGTGGAGTACCAAGTT 1433
Db 1347 CATTTGGCTATTCTGTACATACATGAAAGCAGAGGTTTAAATAATGAAGATATCAAGT 1406
Qy 1434 TGCAGTAAACATGACACACAGAGCTGCAGGTTTGGGAGCTTGGAACTTGGAACTTGAACGGAATTA 1493
Db 1407 AGTCATCGGAATGACACAACTCAATTAATTAATTTATGATGACTTAGAGCTTGCAGCAATTA 1466
Qy 1494 TGAGTCTTACGTGTGTGCTACTCCAGCTGGGGGCGAGCGCACTCCAGCCAGCCCT 1553
Db 1467 TACTTTTCTACATTTAGCATATATGCCAATGGGAGCCAGCCAGATGTCTGACCATGTGAC 1526
Qy 1554 GGTGCATACACTGACGATGTCCCCAG 1580
Db 1527 ACAGAACTCTAGAGGATGACCCAG 1553

RESULT 8

US-09-877-730-19

; Sequence 19, Application US/09877730

; Patent No. 6465632

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Scoville, John

; APPLICANT: Turner, C. Alexander Jr.

; APPLICANT: Friedrich, Glenn

; APPLICANT: Abuin, Alejandro

; APPLICANT: Zambrowicz, Brian

; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. 6465632a1 Human Phosphatases and Polynucleotides Encoding the

; FILE REFERENCE: LEX-0189-USA

; CURRENT APPLICATION NUMBER: US/09/877,730

; PRIOR FILING DATE: 2001-06-08

; PRIOR FILING DATE: 2000-06-09

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 19

; LENGTH: 2724

; TYPE: DNA

; ORGANISM: homo sapiens

US-09-877-730-19

Query Match 3.5%; Score 217; DB 4; Length 2724;

Best Local Similarity 50.3%; Pred. No. 5.1e-46;

Matches 647; Conservative 1; Mismatches 621; Indels 18; Gaps 4;

Qy 300 AGACAGCATGATGAGGAAGCTCTTAGGATCTGGAAGTCACTGAGGCGACCTATTCTGT 359

Db 45 ATACATCAGTGGTGGGAAGCGAGGAGAGCAGTCCGATGAAGGATTTTATCAGTG 104

Qy 360 TCTGGCCACAGCCGCTAGGAGTGTGGCCAGCCAGGTTGCTGTGTTGAAGTTGCCAC 419

Db 105 CTTGGCATGAACAAATATGGAGCAATCTTTAGTCAAAAGGCTCATCTTGGCTTATCAAC 164

Qy 420 ACTCGAAGACTTCTCTCTGCAACCCCGAGTCCAGATTTGTGGAGGAGAAACGGGACACGACG 479
Db 165 TATTTCTGATTTGAAGTCCAGCCAATTTCCACTCAGGTCCACGAAGGTGAGTTGCTCG 224
Qy 480 CTTTGAATGCCACACCAAGGGCTTCCAGCCCCCATCATTAATCTTGGGAAAAGGACCAAGT 539
Db 225 ATTTGATCAAGATTTTATCCACCCCTCTCGAGTCATAACATGAGGATTCATTCGGAC 284
Qy 540 GACCGTGCCTGAG---GAGCCCCGGCTCATCACTCTTCCCAAGTGGCTCTCCACAGATCT 596
Db 285 AACTCTACCTATGACTATGGAAGGATTAAGTCCCTACCAACAGGATTAATTCAGATCTA 344
Qy 597 AGATGTCAGGACAGTATGAGGTCTTACCGCTGCGTGGCCACCAATTCAGCCGCCCA 656
Db 345 TGATGTGACGCAAAAGGATTTCTGAAAATTTATGGTTATTTGCTGCCACTGTAGCCACCG 404
Qy 657 AGATTTACGCGAGGAGGCTCGCTCACTGTGGCCCTCAGAGGCTCTTGTGGAGCTACACG 716
Db 405 AGTAAAGTATGAGGAGGCTCGCTAACTGTG-----ATTCCAGCTTAAGAGTCAAAATC 458
Qy 717 GGGGCAAGGATGTGTCATTTGTGCGAGCCCGCAGAGAACACACCGGTAGTGTCTGGACAGAA 776
Db 459 CTTCCACACACACCAATTTATAGAGGTCCACAGAACATAACAACTCTTTCATCAGAC 518
Qy 777 TGTAGTGTAGTGTGCTCTGCTGACCCCGCTTTTGTGTCTCTGGGTCCGACA 836
Db 519 TGTAGTGTGGAATGATGCGCCACAGAAATCCCAACCAATCATTTCTTGGAGCGCCCT 578
Qy 837 GGNATGGAAGCCAT---CTCCACGATGTCTGTTCTGGGCGGACCAATCTACTCAT 893
Db 579 TGATCAAAATCCATTTGATGTCTTTAATATCTGGGTACTTGGAAATGGTAAATCTCATGAT 638
Qy 894 CGCCAGCGCGAGCTCGGCACCTCTGGAGTCTATGTCTGCGGAGCAACAAAGCCCTCAC 953
Db 639 ATCTGATGTCAGGCTACACATGCTGGAGTATATGTTGTGGGCCACTACCCCTGGCAC 698
Qy 954 GCGTGAATTTGGCCACTGCGGCTGTGAGCTCGAGTCTCGAGTCTGTCGCCAGCCATCTGCA 1013
Db 699 ACGCAACTTTACAGTTGCTATGGCAACTTTAACTGTATTAGTCTCTCTCATTTGTTGA 758
Qy 1014 GGCACCGGAGGCTCTCGGAGCGGGCGGCGAGCACCGCGCTTCTGTGTGCGGCGCTC 1073
Db 759 ATGGCCAGAAAGTTTAAACAAAGGCTCGAGCTGCGACTGTCTGATTTGTGTGTCAGGAGA 818
Qy 1074 CCGGGAGCCACGGCCCGCTGCACTGCGTGCACGACGAGATCCCGTTGCGACCCCAATGG 1133
Db 819 AGGAATCCCTCTCTCCAGATGTCATGTTGAAAATATGAAAGGAGATATATTCGATGG 878
Qy 1134 GCGCGTCAAGGTGACGGCGGTGGCGCAGCTTTGGTCACTCATCTCAGATCGGCTGCAGGA 1193
Db 879 TAGAATTAATAATGTAACAACAGT-----AAATTTGGTAAATTAACAGATTAATTCCTGAAGA 932
Qy 1194 CGCTGGTACTACAGTGTGAGCAGAAAACAGCGGAGGAGTGGCTGTGTCGCTGCGCTGCGCC 1253
Db 933 TGATGCTATTTATCAGTGTGATGAGAAATAGCAAGGATCTATTTTATCTAGAGCCAG 992
Qy 1254 CTTGGCGGTAGTGTGCGGAGGCGGCTGCCCGCGCCGCTCGGCTCAGAGCCAGCCGCC 1313
Db 993 ACTGACTGTAGTGTGAGAGACAGACCCAGTGTCTCCCTAATATGATGATGCTGTAAC 1052
Qy 1314 GCTGAGCAGCTCTCTGTGCTGGCTGGGAGCGGCTGTGAGTGTGACAGCGAGCAAT 1373
Db 1053 CATGTCAAGCTCAGCCATTTCTTTAGCTGGAGAGGAGCCACTTTATATTTTACAGCAAGT 1112
Qy 1374 CATTTGGCTTCTCTTCACTACCAAAAGGAGGAGGAGTGGACAAATGTGAGTACCAAGTT 1433
Db 1113 CATTTGGCTATTTCTGTACACTACATGAAGCAGAAAGGTTTAAATAATGAAGATATCAAGT 1172
Qy 1434 TGCAGTAAAACATGACACACAGAGCTGCGAGTTCGGGAGCTGGAACCCCAACCGAATTA 1493
Db 1173 AGTCATCGGAAATGACACAACTCATTTATATTTATGATGATCTTAGAGCTCTGCCAGCAATTA 1232
Qy 1494 TGAGTTCTAGTGTGGTCTACTCCAGCTGGGGGCGAGCCGAAACCTTCCAGCCAGCCCT 1553

[illegible]

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 21

; LENGTH: 2139

; TYPE: DNA

; ORGANISM: homo sapiens

US-09-877-730-21

Query Match 3.5%; Score 213.8; DB 4; Length 2139;

Best Local Similarity 47.0%; Pred. No. 3.6e-45;

Matches 793; Conservative 0; Mismatches 842; Indels 54; Gaps 2;

```
QY 1158 CGGACGTTGGTCAATCAGATCGGCTGCGAGGCGTGGCTACTACGAGTGGTAGC 1217
DB 60 CAGTAAATGGTAATTAACAGATATTCTCGAAGATGATGCTATTATCAGTGCATGCG 119
QY 1218 AGAAACAGCGCGGAACTGCTGTGCGCTGCGCCCTGCGGTAGTGGTGGCGAGG 1277
DB 120 TGAGAAATAGCCAAAGATCTATTTTATCTAGAGCCAGACTGTATGTATGTGCAAGA 179
QY 1278 GCTGCCAGCGCCCGGACTCGGCTCACAGCCACGCGCTGAGCAGCTCCTCTGTGCTGGT 1337
DB 180 CAGACCCAGTCTCCCTATTAATGTACATGCTGAAACATGTCAAGCTCAGCCATTTCTTTT 239
QY 1338 GGCTGGAGCGGCTGAGTTGCACAGCGAGCAAAATCAATGGCTTCTCTTCACTACCA 1397
DB 240 AGCTGGGAGAGGCCACTTATAATTTCAGACAAAGTCATTGCTTCTGTACACTACAT 299
QY 1398 AAAGGCAAGGGAGTGGCAATGTGGAGTACAGTGTGCGAGTAAACATGACACACAGA 1457
DB 300 GAAAGCAGAAAGGTTAAATTAATGAAGAGTATCAAGTAGTCAATCGGAAATGACACACTCA 359
QY 1458 GCTGCAGTTTCGGGACCTGCAACCCAAACGAGATTATGAGTTCTAGTGGTGGCTACTC 1517
DB 360 TTATATATTATGATGACTTAGAGCTGCGACAAATTAATCTTCTACATTTAGCATATAT 419
QY 1518 CCAGCTGGGCGCCAGCGAACTCTCCAGCCGAGCCCTGGTGATACACTGGAAGATGCC 1577
DB 420 GCCAATGGGAGCCAGCGAGATGCTGACCACTGTGACACAGAACTCTAGAGGATGTTCC 479
QY 1578 CAGGCGAGCAGCCAGCTTACTTGTCCAGCCGCAACCCCTCGGACATCAGGTTGGATG 1637
DB 480 CTTGAGACCTCTCGAAATTAATTTGAACAAGTCGAAGTCCCACTGATATTTCTATCTCTG 539
QY 1638 GCTGCCCTCCCTCCAGCTGAGCAATGACAGGTGCTCAAGTACAAAGTACAGTACGG 1697
DB 540 GCTGCCAATCCAGCCAAATATCGCGGGCCAGTGGTCTGTATGCTGTCTTCTTCCG 599
QY 1698 TTTGGGAAGGAAGATCAGGTTTCTCCACCGAGGTGCTCGAAATGAGACACAACTTAC 1757
DB 600 CTTAAGTACTGAGAAATTCATCCAAAGTTCTGGAGCTCCCGGGACCAACGATGATACCT 659
QY 1758 GTTAAATCTACTCAGCCAAACAAAGTGTACCGAGTCCGATTTTCAGCTGGGACTGGCGC 1817
DB 660 TTTGGAAGGCTGAAACCTGACAGTGTCTACCTGTTTGGATTTACTGCTGCCACCAAGT 719
QY 1818 TGCTATGAGTCCCTTCTCAGTGGATGACGACGACACCTGGTGTGACACACAGAG 1877
DB 720 GGGGCTGGGAGAGTCATGATATGGAATTCACATAGAGCCGCAAGACTCAAGCGGTGAA 779
QY 1878 CCATGTTTCCCTTGTCCCTTCAGAAATTAAGAGGTGAGGGCAAGATGAGTCCCTGTGGT 1937
DB 780 AGCCCTTAAGTCTCCAGATTGCAATTTGGAGCTCTGAACTGTACCACTTTCTGTGAG 839
QY 1938 GTCATGGAGCGGCCCCCTCAACCCCAACAGATCTCTGGATACAACTCTACTGGGAGA 1997
DB 840 GTGGCAGCAAGATGTAGAGGACACAGCTGTCTTTCAGGGGCTTACAAAGCTGTACTACAGGA 899
QY 1998 GGTGGGAACAGAGGAGGAGGAGATGTGTACCGCCCCCAGGGGCTGTGGAGATCAAGC 2057
DB 900 AGAAGGGCAGCAGAGAA----- 917
QY 2058 TTGGGACGTCGGGCCCCGTGGGCTGAAGAGAAAGTGAAGCAGTATGAATGACCCAGTT 2117
```

```
DB 918 -----TGGGCCCATTTTCTTGATACCAAGGACCTACTCTATATCTCTCAGTGGCTT 968
QY 2118 AGTCCCTGGCAGCGCTAGAGGTGAAGCTCTAGTCTTTCAACAAACACGAGGACGGCTTA 2177
DB 969 AGACCCGAGAGAAATATATCTGTGAGTCTCTGCTTACAAACATAGAGAGTGGCTTA 1028
QY 2178 CCGTCTGTGTGGAGGCAAGACGAGAAAGGCGCCAGCGCAGACCTGCTTATCCAGAG 2237
DB 1029 TCAGGCAGATCAGACTGTGAGCAGC---TCCAGGATGCGTGTCTGTTCGTGATCGCATGT 1085
QY 2238 GGGGCCACCGCTGCTCTGCCCCATGTCCACGAGAGTCAACAGCTCACTTCCATTTG 2297
DB 1086 CCGTCTCCACACACACCCACCATCTCTATGCGAAGGCTAAACCTCATCTTCCATCTT 1145
QY 2298 GCTTCGGTGGAGAAAGCCAGACTTTTACACTGTCAAGATTGTCAACTACACTGTAGCTT 2357
DB 1146 CTTGCACTGGAGAGGCTGCAATTCACGCTGCACAAATCATTAATCAACCATCGCTG 1205
QY 2358 CGGCCCTTGGGGCTCAGGAATGCTTCCCTGCTGCTACCTACTATACAGCTCTGGAGAAGA 2417
DB 1206 TAATCTGTGGCTGCGAATGCTTTCTTGGTCTGTACTTCAAAACATCAGAACTCA 1265
QY 2418 CATTCATTTGGCGCTGMAACCAATTTTACCAAGTACGAGTTTGGGTTGCGGTACAGTCCACGG 2477
DB 1266 CATGTTGGTTCAGGCTCTAGAAACCAAAACCAAAATACGAATTTGCGGTTGATTAATGT 1325
QY 2478 AGTGGATATGAGTGGGCCCTTTGGCTCGGTAGAAAGCTCCACCTCGCTGACCGGCC 2537
DB 1326 GATCAGCTTTCCAGTCTCTGGAGCCCTGTAGTCTACCAATCTACTCTTCCAGAGCACC 1385
QY 2538 TTCAACACCTCTCTGACCTCGCTGAGCCCCCTGACACCATCCACGTTGGTTTACA 2597
DB 1386 AGCAGGCCCAACAGTTGGAGTAAAGTGACATTAATAGAGGATGACACTGCCCTGGTTTC 1445
QY 2598 CTGGTGTCCCCCAGCGAGCCCAATGGTGGAGATTGTGGAGTATCTAATTTCTCTACAGCAA 2657
DB 1446 TTGGAACCCCTGATGGCCAGAAACAGTTTGTGACCCGCTATCTATCTTATGTCATC 1505
QY 2658 CAACACACCCAGCCGGAACACACAGTGGACACTGCTCACACAGAGGGAACATCTTCAG 2717
DB 1506 TAGAAGGCTGTGATTTGAGGAGTGGCAGGCTTACACCGTGAAGGGGCAATAACCAT 1565
QY 2718 TGCAAGGTCCATGGCTGAGAGTGACACTCGGTATTTTCTTCAAGATGGGAGCCGCGAC 2777
DB 1566 GCGTTTGTAGAAACCTTGTAGCAGGAAATGTGTACATTTGTCAAGATATCTGCATCCAA 1625
QY 2778 AGAGTGGGCGCTGGGCCCTTTTCCGCTTGGCAGGATGTGATTAATCTCTGCAAGAGACAT 2837
DB 1626 TGAGTGGGAGAGGAGCCCTTTTCAAAATTTCTGTGGAGCTGGCAGTACTTCCAAAGGAAAC 1685
QY 2838 CTCAGACTC 2846
DB 1686 CTCGATC 1694
```

RESULT 11

US-09-877-730-27

; Sequence 27, Application US/09877730

; Patent No. 6465632

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Scoville, John

; APPLICANT: Turner, C. Alexander Jr.

; APPLICANT: Friedrich, Glenn

; APPLICANT: Abuin, Alejandro

; APPLICANT: Zambrowicz, Brian

; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding

; FILE REFERENCE: LEX-0189-USA

; CURRENT APPLICATION NUMBER: US/09/877,730

; PRIOR FILING DATE: 2001-06-08

; PRIOR APPLICATION NUMBER: US 60/210,607

; PRIOR FILING DATE: 2000-06-09

NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FactSeq For Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 2382
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-27

Query Match 3.5%; Score 213.8; DB 4; Length 2382;
Best Local Similarity 47.0%; Pred. No. 3.8e-45;
Matches 793; Conservative 0; Mismatches 842; Indels 54; Gaps 2;

```
QY 1158 CGGCAGCTGGTCATCCTCAGATCGGCTGCGAGACGCTGGCTACTACCACTGGTGGTACG 1217
DB 60 CAGTAAATGGTAAATTAACAGATTAATTCCTGAAGATGATGCTATTTATCACTGCGATGCG 119
QY 1218 AGAAACAGCGCGGAACTGCTGTGCGCTGCGCCCTGGCGGTAGTGGTGGCGGAGG 1277
DB 120 TGAGATAGCCNAGATCTATTTATCTAGAGCCAGACTGACTGTAGTGTGTCAGAGA 179
QY 1278 GCTGCCAGCGCCCGACTCGGGTCAAGCCACGCGCTGAGCAGCTCTCTGTGTGGT 1337
DB 180 CAGACCCAGTCTCCCTAATATGATATGCTGAACCATGTCAAGCTCAGCCATTTCTTTT 239
QY 1338 GGCCTGGGAGCGGCTGTGATGTCAGCGAGCAATATGCTTCTCTCTACTACCA 1397
DB 240 AGCCTGGGAGGCGCCATTTAATTTCAAGCAAGTCAATGCTTATTTCTGTACACTACAT 299
QY 1398 AAGGCAAGGGAGTGGCAATGTGAGTACCAGTTTCAGTAAACATGACACACAGA 1457
DB 300 GAAACAGAGGTTTAAATATGAAGATATCAAGTATGCTATCGGAATGACAACTCA 359
QY 1458 GCTGAGGTTCGGGACCTGGAACCCCAACAGGATTTAGTGTTCAGTGGTGGCCTACTC 1517
DB 360 TTATATTATGATGACTTAGAGCTGCGCAGCAATTTACTTTCTACATTTAGCATATAT 419
QY 1518 CCAGTGGGGGCGCCGAGACCTCCAGCCGAGCCCTGGTGCATACACTGAGCATGTCCC 1577
DB 420 GCAATGGGAGCGACCGCAGATGTCTGACCATGTGACACAGAATACTCTAGAGGATGTTC 479
QY 1578 CAGGCGACGACCCAGCTTACTGTGCGAGCCCAACCCCTCGGACATCAGGGTGGCATG 1637
DB 480 CTGAGACTCTCTGAAATAGTTTGAAGTGTGAAGTCCCACTGATATTTCTCATCTCCTG 539
QY 1638 GCTGCCCTGCGCTCCAGCTGAGCAATGGACAGGTGCTGAAGTACAGATAGAGTACGG 1697
DB 540 GCTGCCAATCCAGCCAAATATCGCGGGGCAAGTGGTGTGATCGCTTCTTTCG 599
QY 1698 TTGGGGAGGAAGATCAGTTTCTCCACGAGGTGCTGGAATGAGACACAACTTAC 1757
DB 600 CTAAGTACTGAGAAATCAATCCAAGTTCTGAGCTCCCGGGACCAAGCATGAGTACCT 659
QY 1758 GTTAAACTCACTTCAGCCAAAGTGTACCGAGTCCGGATTTCAAGTGGCACTGGCGC 1817
DB 660 TTGGAAAGGCTGAAACCTGACAGTGTCTACCTGGTTCCGATTTACTGTGCCACAGAGT 719
QY 1818 TGGCTATGAGTCTCTCAGTGTGATGCGACAGGACACTGGTGTGCACACAGAG 1877
DB 720 GGGGCTGGGAGTGTATCAGTATGAGCTTCAATAGGAGCGCCAAAGCTACAGGCTGAA 779
QY 1878 CCATGTTCCCTTGGCCCTGCGAATTTGAAGGTGAGGCAAGATGAGAGTCCCTGGTGGT 1937
DB 780 AGCCCTAAGTCTCCAGAGTTGCAATTTGGAGCTCTGAACCTGTACCACTTTCTGTGAG 839
QY 1938 GTCATGGAGCGGCGCCCTGACCCCAACAGATCTCTGGATACAACTCTACTGGGGAGA 1997
DB 840 GTGGCAGCAAGATGTAGAGGACACAGCTGCTATTCCAGGGCTACAGCTGTACTACAAGA 899
QY 1998 GTTGGAAACAGAGGAGGAGGAGATGGTGAACCGCCCCCGAGGGGTCTGTGGAGATCAAGC 2057
DB 900 AAGAGGCGAGCAGGAGAA----- 917
QY 2058 TTGGGACGTCGGGCGGCTGGGCTGGAAGAAAGTGAAGCATGATGAACCTGACCCAGTT 2117
```

```
DB 918 -----TGGGCCCATTTTCTTGGATACCAAGGACCTACTCTATACTCTCACTGCGCTT 968
QY 2118 AGTCCCTGGCAGGCGCGTACGAGGTGAAGCTCGTAGCTTTTCAACAAACACAGGAGCGGCTA 2177
DB 969 AGACCCAGAGAAATATCATGTGAGACTCTCTGCTTTACAAACATAGACGATGGCTA 1028
QY 2178 CGCTGTGTGTGGAAGGCAAGACGGAGAGGCGCCCAACGCGACAGCTGCTTCACAGAG 2237
DB 1029 TCAGGCAGATCAGACTGTGTGAGCAC---TCCAGGATGGGTGTCTGTCTGTGATCGCATGGT 1085
QY 2238 GGGGCCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2297
DB 1086 CCCTCTCCACACACACCCACCATCTCTATGGAAGGCTTAACACCTCATCTTCCATCTT 1145
QY 2298 GCTTTCGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2357
DB 1146 CCGTCACTGGAGGAGGCTGCAATTCACCGCTGCACAATCATTAACCTACACCATCCGCTG 1205
QY 2358 CGGCCCCCTGGGGGCTCAGGAATGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2417
DB 1206 TAATCTGTGTGGCTGCGAGAAATGCTTCTTGGTCTGTGTGCTGCTGCTGCTGCTGCTGCTG 1265
QY 2418 CATTTCTCATTTGGCGGCTGAAACCAATTTACCAAGTACGAGTTTGGGCTACAGTCCCGG 2477
DB 1266 CATGTTGGTTCAAGGCTTAGAACCACAAACACCAATACCAATTTTGGCCGTTGATTTACATG 1325
QY 2478 AGTGGATATGATGGGCGCTTGGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2537
DB 1326 GGATCAGCTTTCAGTCTCTGGAGCCCTGTAGTCTACATTTCTACTCTTCCGAGAGCACC 1385
QY 2538 TTCAACACCTCTCTGACCTGCGCTGAGCCCTGAGCCCTGACACCATTCACCGTTCCGTTTACA 2597
DB 1386 AGCAGGCCACACAGTTGGAGTAAAGTGAATTAATAGAGGATGACACTGCGCTGGTTTC 1445
QY 2598 CTGCTGTCTCCCGCAGGAGCCCAATGGTGAAGTTGTGAGATATCTAATTTCTCTACAGCA 2657
DB 1446 TTGAAACCCCTGATGGCCGAGAAACAGTTGTGACCGCTTATCTATCTATATGCTATC 1505
QY 2658 CAACACACCCAGCCGCAACACACAGTGCACACTGCTCACACAGAGGAAACATCTTCAG 2717
DB 1506 TAGGAGGCTGGATTTGAGAGAGTGGCAGGCTTTACACCGTGAAGGCAATTAACCAT 1565
QY 2718 TGCAAGAGTCCATGCGCTTAGAGGTGACACTCGGTATTTCTTCAAGATGGAGCGCCGAC 2777
DB 1566 GCTTTGTCTAGAAAACCTTGGTAGCAGGAATGTGTACATTTGTCAAGATATCTGCAATCAA 1625
QY 2778 AGAGTGGGGCTGGGCGCTTTTCCCGCTTGCAGGATGTGATTTACTCTGCAAGAGACATT 2837
DB 1626 TGAGTGGGAGAGGAGCCCTTTTCAAAATTTCTGTGAGCTGGCAGTACTTCCAAAGGAAC 1685
QY 2838 CTCAGACTC 2846
DB 1686 CTCTGAATC 1694
```

RESULT 12

US-09-877-730-23

; Sequence 23, Application US/09877730

; Patent No. 6465632

; GENERAL INFORMATION:

; APPLICANT: Scoville, John

; APPLICANT: Walker, D. Wade

; APPLICANT: Turner, C. Alexander Jr.

; APPLICANT: Friedrich, Glenn

; APPLICANT: Abuin, Alejandro

; APPLICANT: Zambrowicz, Brian

; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding

; FILE REFERENCE: LEX-0189-USA

; CURRENT APPLICATION NUMBER: US/09/877,730

; PRIORITY FILING DATE: 2001-06-08

; PRIOR APPLICATION NUMBER: US 60/210,607

;; PRIOR FILING DATE: 2000-06-09
;; NUMBER OF SEQ ID NOS: 31
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 23
;; LENGTH: 1875
;; TYPE: DNA
;; ORGANISM: homo sapiens
US-09-877-730-23

Query Match 2.4%; Score 151.2; DB 4; Length 1875;
Best Local Similarity 45.8%; Pred. No. 5.4e-29;
Matches 662; Conservative 0; Mismatches 728; Indels 54; Gaps 2;

```
QY 1403 CAAAGGAGTGGCAATGTGGAGTACCAAGTTTCAGTAAACATGACACACAGAGCTGC 1462
DB 41 CAACAGGTTAAATATAGAGAGTATCAAGTAGTCAATCGGMAATGACACATCAATATA 100

QY 1463 AGGTCGGGACCTGGAAACCAACACCGGATATAGAGTTCTAGCGTGGCTACTCCAGC 1522
DB 101 TTATTGATGACTTAGAGCTGCCAGCAATATATATCTTTCTACATTGTAGCATATATGCCAA 160

QY 1523 TGGGGCCAGCCGAACTCCAGCCAGCCCTGGTGCATACATCGAGGATGCCCCAGG 1582
DB 161 TGGAGCCAGCCAGATGTGTGACCATGTGACACAGAAATCTCTAGAGGATGTTCCCTGA 220

QY 1583 CAGCACCAGCTTACCTTCCAGGCCCAACCCCTCGGACATCAGGCTGGCATGGCTGC 1642
DB 221 GACCTCTGAAATAGTTGACAAAGTGAAGTCCCACTGATATCTCATCTCTGGCTGC 280

QY 1643 CCTGGCCCTCCAGCTGAGCAATGGAAGGTGCTGGAAGTACAAAGATAGAGTAGCGTTGG 1702
DB 281 CAATCCAGCCAAATATCGCGGGCCAGTGGTGTGTATCGTCTTCTTCCGCTAA 340

QY 1703 GGAAGGAGATCAGTTTCTCCAGGAGTGGCTGGAAATGAGACACAACTTACCTTAA 1762
DB 341 GTACTGAGAAATCAATCCAAAGTTCTGGAGCTCCCGGGACCAACGCAATGAGTACCTTTGG 400

QY 1763 ACTCAGTTCCAGCAACAAAGTGTACCGAGTCCGGATTTTCAGCTGGCACTGGCGTGGCT 1822
DB 401 AAGGCTGAACTGACAGTGTCTACTCTGTTGGATTTGATTTGCTGCCACGAGTGGGC 460

QY 1823 ATGAGTCCCTTCTCAGTGGATGACGACAGGACACCTGGTGTGCAACACAGAGGCAATG 1882
DB 461 TGGGAGAGTATCAGTATGAGCTTCAATAGGAGCCCAAGCTACAGGCGTGAAGGCC 520

QY 1883 TTCCCTTTGGCCCTGCAAGATTTGAAGTGAAGGCAAGATGGAGTCCCTGGTGTCTCAT 1942
DB 521 CTAAGTCTCAGAGTTGCATTTGGAGCTCTGAACTGTACCAACATTTCTGTGAGGTGGC 580

QY 1943 GGCAGCGCCCTCCAGCCAGATCTCTGGATACAACTCTACTGGGGAGAGGTGG 2002
DB 581 AGCAAGATGTAGAGGACACAGCTGCTATTTCAGGGCTACAGCTGTACTACAAGGAAGAG 640

QY 2003 GAACAGAGGAGGAGGAGATGGTGAACCGCCCGGAGTGGTGGAGATCAAGCTTGGG 2062
DB 641 GGCAGCAGGAGAA-----653

QY 2063 ACCTGGGCGCGTGGCTGAGAGAAAGTGAAGCAGTATGAATGACCCAGCTTAGTCC 2122
DB 654 -----TGGGCGCAATTTCTTGGATACCAAGGACCTACTCTATCTCTCAGTGGCTTAGAC 709

QY 2123 CTGGCAGCGCTGACAGCTGAAGCTCGTAGCTTTCAACAAACAGGAGCGGTACGCTG 2182
DB 710 CCAGAAGAAATATCATGTGAGACTCTGCTGCTTCAACAAACATAGACGATGGCTATCAGG 769

QY 2183 CTGTGTGGAAGGGCAAGAGGAGGCGCCAGCCAGACCTGCTTATCCAGAGGGGC 2242
DB 770 CAGATCAGAGTGTGAGCAGC---TCCAGGATGGTGTCTGTTGATGCGATGCTGCCCTC 826

QY 2243 CACGGTCTCTTGGCCATGTGCCAGAGTCAACAGCTCCACTTCCATTGGCTTC 2302
DB 827 CTCACACCAACCCAGCATCTCTATGCGAAGGCTAACACCTCATCTTCCATCTTCCCTGC 886
```

```
QY 2303 GGTGGAAGAGCCAGACTTTTACCAGTGTCAAGATGTCAACTACACTGTAGCGTTGGCC 2362
DB 887 ACTGGAGGAGGCGCTGCAATTCACCGCTGCACAAATCAITAACTATACCACTCCGCTGTAATC 946

QY 2363 CCTGGGGCTCAGGAATGCTTCCCTGGTCACTACTATACAGCTCTGGAGAGACATTC 2422
DB 947 CTGTTGGCTTCAGAAATGCTTTCTTGGTCTGTCTTCAAAATCAAGAACTCATATGT 1006

QY 2423 TCATTTGGGGCTGAAACCAATTTTACCAAGTACGAGTTTGGGTGACAGTCCCAAGGAGTGG 2482
DB 1007 TGGTTCAAGGTCTAGAACCAACACCAATACGAATTTGGCGTTGATTTACATGTGGATC 1066

QY 2483 ATATGGATGGGCCCTTTTGGCTCCGTCGTGAGAACGCTCCACCCCTGCTGACCGGCTTCAA 2542
DB 1067 AGCTTTCCAGTCTCTTGGAGCCCTGTAGTCTACATTTCTACTCTTCCAGAAGCACCAAG 1126

QY 2543 CACCTCTCTTCTGACCTGGCTGGCCCTGACCCCTGACACCATCCACCGTTGGTTTACACTGT 2602
DB 1127 GCCCAACCAAGTTGGAATGAAAGTGACATTAATAGAGATGACACTCCCTGGTTTCTTGA 1186

QY 2603 GTCCCCCAGCAGGCGCAATGTGTGAGATTTGTGAGATATCTAATTTCTACAGCAACAACC 2662
DB 1187 AACCCCTGATGGCCAGAACAGTTGTGACCCGCTATACTATCTTATATGATCTAGGA 1246

QY 2663 ACACCCAGCCGAAACACAGTGAACACTGCTCCACACAGAGGGAACATCTTCACTGAG 2722
DB 1247 AGGCTGTGATTCAGGAGAGTGGCAGGCTTACACCGTGAAGGGCAATAACCATGGCTT 1306

QY 2723 AGTCCATGGCTTAGAGAGTGAACACTCGGTATTTCTCAAGATGGGAGCCGACAGAGG 2782
DB 1307 TGTAGAAACACTTGTGTAGCAGGAATGTGTACATTTGCAAGATATCTGCATCCAATGAGG 1366

QY 2783 TGGGGCTGGGCGCTTTTCCCGCTTTCAGAGATGTGATTTACTCTGCAAGAGACATTTCTCAG 2842
DB 1367 TGGGAGAGGAGCCCTTTTCAATTTCTGTGGAGCTGGCAGTACTTCCAAAGGAACCTCTG 1426

QY 2843 ACTC 2846
DB 1427 AATC 1430
```

RESULT 13

US-09-877-730-13
; Sequence 13, Application US/09877730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding
; FILE REFERENCE: LEX-0189-USA
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 909
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-13

Query Match 1.9%; Score 119; DB 4; Length 909;
Best Local Similarity 48.7%; Pred. No. 8e-21;
Matches 416; Conservative 1; Mismatches 426; Indels 12; Gaps 3;
QY 300 AGACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAGGTGACCTGAGGGGAGCTATTCCTG 359

Db 1113 TAGAATTAAATGTA 1127

RESULT 15

US-09-877-730-25
; Sequence 25, Application US/09877730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding the
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 1644
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-25

Query Match 1.9%; Score 117; DB 4; Length 1644;
Best Local Similarity 48.4%; Pred. No. 3.6e-20;
Matches 355; Conservative 0; Mismatches 375; Indels 3; Gaps 1;
QY 2114 AGTTAGTCCCTGGCAGGCGCTAGAGGTGAAGCTCTAGCTTTCAACAACACGAGGAG 2173
Db 470 AGGATGACCCAGAGAAATATCATGTGAGACTCTCTGCTTACAAACATAGACGATG 529
QY 2174 GCTACGCTGTGTGTGGAGGGCAAGACGGAGAAGCGCCACGACGAGACCTGCCTATCC 2233
Db 530 GCTATCAGGAGATCAGATCTGTGAGC---TCAGGATGCGTCTGTTGATCGCA 586
QY 2234 AGAGGGGGCCACCGCTCTCTCCCATGTCCACGAGAGTCAAAACAGCTCCACCTTCCA 2293
Db 587 TGGTCCCTCTCTCCACGACACCCACCATCTCTATGGAAGGCTAACACCTCATCTTCCA 646
QY 2294 TTTGGCTTCGGTGGAGAGAGCCAGACTTTACCACTGTCAAGATTGTCAACTACACTGTAC 2353
Db 647 TCTTCTGCACTGGAGAGGCGCTGCATTTACCGCTGCACAAATCATTAACCTACCATCC 706
QY 2354 GCTTCGGCCCTGGGGGCTCAGGAATGCTTCCCTGCTCACCCTACTATACGAGCTCTGGAG 2413
Db 707 GCTGTAACTCTGTGGCTGAGAAATGCTTCTTTGGTCTGTACCTTCAAAACATCAGAA 766
QY 2414 AAGACATTTCTCATTTGGCGGCTTGAACCAATTTACCAAGTACGAGTTTGGGTACAGTCCC 2473
Db 767 CTCACATTTGGTTCAAGGTCTAGAACCAACCAAAATACGAATTTGCGTTGATTTAC 826
QY 2474 ACGGAGTGATATGAGATGGGCCCTTTGGGCTTCGCTGTAGAAACGCTCCACCCCTGACC 2533
Db 827 ATGTGGATCAGCTTTTCAGTCTTGGAGCGCTGTAGTCTACCATTTCTACTTCTCCAGAAG 886
QY 2534 GGCCTTTCAACACCTCTCTGACCTGGCTGAGCCCTCGACCATCCACCGTTGGT 2593
Db 887 CACGAGAGGCCACCATTTGGAGTAAAGTGAATTAATAGAGGATGACACTGCCCCGG 946
QY 2594 TACACTGTGTTCCTCCCAACGAGCCCAATGGTGGAGATTGTGGAGTATCTAAATTTCTTACA 2653
Db 947 TTTCTTGGAAACCTCTGTATGGCCAGAAACAGTTGTGACCCGCTATCTATCTATATG 1006
QY 2654 GCAACACCAACACCCAGCCCGAAACACGATGGACACTGTCTCACACAGAGGGAAACATCT 2713
Db 1007 CATCTAGGAAGGCGCTGATTTGAGGAGAGTGGCAGGTCTTTACACCGTGAAGGGCAATAA 1066

QY 2714 TCAGTGCAGAGGTCCATGGCCTTAGAGAGTGACACTCGGTATTTCTTCAAGATGGAGCCC 2773
Db 1067 CCATGGCTTTGCTTAGAABAACTTTGGTAGCAGGAAATGTGTACATTGTCAAGATATCTGCAT 1126
QY 2774 GCACAGAGGTGGGCGCTGGGCCCTTTTCCCGCTTGCAGGATGTGATTTACTCTGCAAGAGA 2833
Db 1127 CCAATGAGGTGGGAGAGGAGCCCTTTTCAAAATTTCTGTGAGCTGGCAGTACTTTCCAAAGG 1186
QY 2834 CATTTCTCAGACTC 2846
Db 1187 AAACCTCTGAATC 1199

Search completed: October 15, 2003, 04:46:31
Job time : 383.754 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 14, 2003, 06:11:05 ; Search time 7190 Seconds
(without alignments)
7123.623 Million cell updates/sec

Title: US-09-754-997A-2

Perfect score: 6604

Sequence: 1 MARADTCRGLLVLTFCLLSA.....CPLTVSPSLPRAPVSSAQP 1252

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09754997/runat_14102003_061100_9706/app_query.fasta_1.1415
-DB=GenEmbl -QFWT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOFCL=0 -LOPEXT=0
-UNITS=dbs -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USBR=US09754997@cgn_1_1_4823@runat_14102003_061100_9706 -NCPU=3
-NO MAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DBLOCK=6 -DELEXT=7

Database :

GenEmbl:

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_cm:*
21: em_ov:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_man:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6604	100.0	3756	6	AX191303 Sequence
2	6604	100.0	6176	6	AX191259 Sequence
3	6604	100.0	6176	10	AF176694 Mus muscu
4	6532	98.9	6219	10	AB052620 Mus muscu
5	6521.5	98.8	6222	10	AB052621 Mus muscu
6	6351.5	96.2	6301	10	AK122535 Mus muscu
7	5746	87.0	6485	9	AB052622 Homo sapi
8	5610	84.9	3741	6	AX18406 Sequence
9	4903	74.2	2796	6	AX191261 Sequence
10	4361.5	66.0	5549	9	AB046848 Homo sapi
11	3125.5	47.3	158055	2	AC110235 Mus muscu
12	3064.5	46.4	224227	2	AC106509 Rattus no
13	2557	38.7	176282	2	AC068507 Homo sapi
14	2557	38.7	198295	9	AC105129 Homo sapi
15	2466.5	37.3	116580	10	AC112162 Mus Muscu
16	2321	35.1	115135	10	AC112161 Mus Muscu
17	1747.5	26.5	3453	6	AR237556 Sequence
18	1746.5	26.4	3210	6	AR237553 Sequence
19	1746.5	26.4	3874	6	AR237568 Sequence
20	1691.5	25.6	2976	6	AR237558 Sequence
21	1691.5	25.6	3219	6	AR237561 Sequence
22	1472	22.3	825	6	AX191263 Sequence
23	1384	21.0	2715	6	AR237555 Sequence
24	1384	21.0	2958	6	AR237557 Sequence
25	1329	20.1	2481	6	AR237560 Sequence
26	1329	20.1	2724	6	AR237562 Sequence
27	1144	17.3	2139	6	AR237563 Sequence
28	1144	17.3	2382	6	AR237566 Sequence
29	1070.5	16.2	2967	9	BC042054 Homo sapi
30	1067	16.2	2700	6	AX457857 Sequence
31	1046	15.8	3198	10	BC053057 Mus muscu
32	1031.5	15.6	3146	10	AF026465 Mus muscu
33	978	14.8	4784	5	GU07644 Gallus gall
34	953.5	14.4	4631	9	HSU72391 Human neoge
35	951.5	14.4	5297	6	AX587800 Sequence
36	951.5	14.4	5297	9	HSU61282 Human neoge
37	947	14.3	4641	10	RNU68726 Rattus norv
38	945.5	14.3	5414	10	MMNEOGEN Y09535 M.musculus
39	943	14.3	4344	6	AX054981 Sequence
40	943	14.3	4608	6	AR153583 Sequence
41	943	14.3	4608	9	HSDCCG X76132 H.sapiens D
42	940.5	14.2	4517	10	MMDCCTSG X85788 M.musculus
43	932.5	14.1	4846	5	XLU10986 Xenopus lae
44	931	14.1	4924	10	RNU68725 Rattus norv
45	921.5	14.0	2729	9	AK095529 Homo sapi

ALIGNMENTS

RESULT 1

AX191303
 LOCUS AX191303 3756 bp DNA linear PAT 15-AUG-2001
 DEFINITION Sequence 45 from Patent WO0149714.
 ACCESSION AX191303
 VERSION AX191303.1 GI:15209554
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Salbaum, J.M.
 Nope polypeptides, encoding nucleic acids and methods of use
 TITLE Patent: WO 0149714-A 45 12-JUL-2001;
 JOURNAL Neurosciences Research Foundation Inc. (US)
 FEATURES
 source 1..3756
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 BASE COUNT 774 a 1167 c 1105 g 710 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.74e-260 Length: 3756
 Score: 6604.00 Matches: 1252
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-754-997A-2 (1-1252) x AX191303 (1-3756)
 QY 1 MetAlaArgAlaAspThrGlyArgGlyLeuLeuValLeuThrPheCysLeuLeuSerAla 20
 DB 1 ATGGCGCGGCGGACACGCGCGCGGCTCTGGTGTGACCTTCTGCGCTGCTCGCG 60
 QY 21 ArgGlyLeuLeuProLeuProGlnGluThrValLysLeuSerCysAspGluGlyPro 40
 DB 61 CGCGGGAGGCTGCANTTGCCTCCAGGAGACACTGTCTAGCTGAGCTGTGATGAGGACCC 120
 QY 41 LeuGlnValLeuLeuGlyProGluGlnAlaValValLeuAspCysThrLeuGlyAlaThr 60
 DB 121 CTGCAAGTGATCTCGGCGCTGAGCAGGCTGTGTCTGAGCTGCACCTTTGGGGGTACA 180
 QY 61 AlaAlaGlyProProThrArgValThrTrpSerLysAspGlyAspThrValLeuGluHis 80
 DB 181 GCTGTGGGCTCCGACCGAGGTGACATGGAGCAAGGATGGAGACACTGTACTAGAGCAT 240
 QY 81 GluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGluGlnGlu 100
 DB 241 GAGAACTGTGACCTGTCTACCAATAGGCTCCCTGTGGCTGTCTACCCCTAGAGCAAGAA 300
 QY 101 AspSerAspAspGluGluAlaLeuArgIleTrpLysValThrGluGlySerTrpSerCys 120
 DB 301 GACAGCGATGATGAGGAGCTCTTAGGATCTCGAAGGTCACTCAGGGGAGCTATTCTGT 360
 QY 121 LeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValValLysLeuAlaThr 140
 DB 361 CTGGCCACACAGCGCTAGGAGTGTGGCCAGCGAGGTTGCTGTGGTCAAGCTTGGCCACA 420
 QY 141 LeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThrAlaArg 160
 DB 421 CTGAAAGACTTCTCTGCAACCCGAGTCCAGATTGTGGAGAGAACCGGACAGCACGC 480
 QY 161 PheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAspGlnVal 180
 DB 481 TTGTGAATGCCACACCAAGGCGCTTCCAGCCCTCATCTACTTGGGAAAGGACAGGTG 540
 QY 181 ThrValProGluGluProArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuAsp 200
 DB 541 ACCGTGCTGAGGAGCGCGGCTCATCACTCTTCCCAAGTGGCTCTCTCCAGATCCTAGAT 600

QY 201 ValGlnAspSerAspAlaGlySerTrpArgCysValAlaThrAsnSerAlaArgGlnArg 220
 DB 601 GTCCAGGACAGTGATGCAAGGCTCTTACCGTGGTGGCCACCAATTACGCCGCCCAACA 660
 QY 221 PheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGly 240
 DB 661 TTCCAGCCAGGAGGCGCTCGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTACCAAGGG 720
 QY 241 GlnAspValValIleValAlaAlaProGluAsnThrThrValValSerGlyGlnAsnVal 260
 DB 721 CAGGATGTGGTCTATTGTGGCAGCCCGAGAGAACACACCGTAGTGTCTGGACAGATGTA 780
 QY 261 ValMetGluCysValAlaSerAlaAspProThrProPheValSerTrpValArgGlnAsp 280
 DB 781 GTGATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 840
 QY 281 GlyLysProIleSerThrAspValIleValLeuGlyArgThrAsnLeuLeuLeuAsp 300
 DB 841 GGAAGAGCTTATCTCCAGCGATGTCTGCTTGGGCGGACCAATCTACTATCCATCCACG 900
 QY 301 AlaGlnProArgHisSerGlyValTrpValCysArgAlaAsnLysProLeuThrArgAsp 320
 DB 901 GCGCAGCTCGGCACTCTGGAGTCTATGTCTCCGAGCGCACAGCCCTCTCAGCGTGAC 960
 QY 321 PheAlaThrAlaAlaAlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaPro 340
 DB 961 TTTCGCCACTGCGGCTGTGAGCTCCGAGTGTCTGCTGCCCGCAGCATCTCGCAGGACCC 1020
 QY 341 GluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGlu 360
 DB 1021 GAGGCGCTCTCGCGAGCGCGGCGGACCGCGCGCTTCTGTGGCGGCGCTCGCGGGAG 1080
 QY 361 ProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgVal 380
 DB 1081 CCAGCGCCCGCTGCACTGGCTGCAGCGGAGTCCCTGCGACCCCAATGAGCGCGCTC 1140
 QY 381 LysValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGly 400
 DB 1141 AAGGTGAGGCGGCTGGCGGAGCTTGTCTCATCTCAGATCGGCTCGCAGGACGCTGGC 1200
 QY 401 TyrTrpGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaAlaProLeuAla 420
 DB 1201 TACTACAGTGTGTAGCAGAAACAGCGGGAACTGCGCTGTGCGCTGTGCGCTCGCGCG 1260
 QY 421 ValValArgGluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSer 440
 DB 1261 GTAGTGTGCGGAGGCGCTGCCAGCGCCCGACTCGGCTCAGAGCCAGCGCTGAGC 1320
 QY 441 SerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGluGlnIleLeuGly 460
 DB 1321 AGCTCTCTGTGCTGGTGGCTGGGAGCGGCTGAGTTGACAGCGAGCAATCATTTGCG 1380
 QY 461 PheSerLeuHisTrpGlnLysAlaArgGlyValAspAsnValGluTrpGlnPheAlaVal 480
 DB 1381 TTCTCTCTTCACTACCAAAAGGCAAGGGAGTGGACAAATGTGGAGTACCAAGTTTGCAGTA 1440
 QY 481 AsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTrpGluPhe 500
 DB 1441 AACAAATGACACACAGAGCTGAGGTTCGGGACCTGGAAACCCCAACACCGATTATGATTC 1500
 QY 501 TyrValValAlaTrpSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHis 520
 DB 1501 TAGTGTGGCTTACTCCAGCTGGGGGCGAGCGCACTCCAGCCAGCCCTGTGGTGCAT 1560
 QY 521 ThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSerProAsnSer 540
 DB 1561 ACACTGGACGATGTCCCGAGCGCAGCACCCAGCTTACCTTGTCTCCAGCGCCCAACCCCTCG 1620
 QY 541 AspIleArgValAlaTrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLys 560
 DB 1621 GACATCAGGTTGGCATGGCTGGCCCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAG 1680
 QY 561 TyrLysIleGluTrpGlyLeuGlyLysGluAspGlnValPheSerThrGluValProGly 580

Db 1681 TACAGATAGAGTACGGTTTGGGAAAGGAAGATCAGGTTTCTCCACCGAGGTGCCCTGA 1740
Qy 581 AenGluThrGlnLeuThrLeuAenSerLeuGlnProAenLysValTyrArgValArgile 600
Db 1741 AATGAGACACAACTTAGCTTAACTCACTTCAGGCCAAACAAGGTGTACCGAGTCCGAT 1800
Qy 601 SerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrPro 620
Db 1801 TCAGTGGCACTGGCGCTGGCTATGGAGTCCCTTCTCAGTGGATGCAGCAAGGACACCT 1860
Qy 621 GlyValHisGlnSerHisValProPheAlaProAlaGluLeuLysValArgAlaLys 640
Db 1861 GGTGTGCACAAACAGAGCCATGTTCCCTTTCCTGCGCTGAGAAATGAAAGGTGAGGCAAG 1920
Qy 641 MetGluSerLeuValSerTrpGlnProProHisProGlnHisSerGlyTyr 660
Db 1921 ATGGAGTCCCTGGTGGTGTATGGACGCGCCCTCACCCACCCAGATCTCTGGATAC 1980
Qy 661 LysLeuTyrTrpGlyGluValGlyThrGluGluAlaAseGlyAspArgProProGly 680
Db 1981 AAACCTACTGGGAGAGGTGGGAACAGAGGAGGAGGAGATGGTGCACCGCCCGCAGGG 2040
Qy 681 GlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLysLysValLysGln 700
Db 2041 GGTGTGGAGATCAAGCTTGGGACGTGGGCGCTGGGCGCTGGAAGAAAGTGAAGCAG 2100
Qy 701 TyrGluLeuThrGlnLeuValProGlyArgProTyrGluValLysLeuValAlaPheAen 720
Db 2101 TATGACTGACCAAGTATGCTTGGCAGCGCGCTACGAGGTGAAGCTCTGATCTTTCAAC 2160
Qy 721 LysHisGluAseGlyTyrAlaAlaValTrpLysGlyLysThrGluLysAlaProThrPro 740
Db 2161 AAACACAGAGGACGCTACGCTGCTGTGTGAAGGCGCAAGACGAGAGGCGCCACGCCA 2220
Qy 741 AspLeuProLleGlnArgGlyProProLeuProProAlaHisValHisAlaGluSerAen 760
Db 2221 GACCTGGCTATTCAGAGGGGCGCCACCGCTGCTCTGCGCATGTCCACGAGAGTCAAC 2280
Qy 761 SerSerThrSerLleTrpLeuArgTrpLysLysProAspPheThrThrValLysileVal 780
Db 2281 AGCTCCACTTCATTTGGCTTGGTGGAGAGGCGAGCTTTACACTGTCTCAAGATTGTC 2340
Qy 781 AsnTyrThrValArgPheGlyProTrpGlyLeuArgAenAlaSerLeuValThrTyrTyr 800
Db 2341 AACTACACTGTACGCTTGGCCCTCGGGGCTCAGGAATGCTTCCCTGGTCACTACTAT 2400
Qy 801 ThrSerSerGlyGluAseLleLeuLleGlyLysLysProPheThrLysTyrGluPhe 820
Db 2401 ACCAGCTCTGGAGAAGACATTTCTATTGGCGGCTGAAACCATTTTACCAAGTACGAGTTT 2460
Qy 821 AlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSer 840
Db 2461 GCGGTACAGTCCCAAGAGTGGATGGATGGGCGCTTGGCTCGGTGAGAGAGCTCC 2520
Qy 841 ThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrPro 860
Db 2521 ACCCTGCTGACCGGCTTCAACACTCTTCTGACCTGGCGCTGAGCGCCCTGCACCA 2580
Qy 861 SerThrValArgLeuHisTrpCysProProThrGluProAenGlyGluLysValGluTyr 880
Db 2581 TCCACCGTTTCGTTTACACTGGTGTCCCCCAGCGAGGCCAATGGTGGATTTGTGGATAT 2640
Qy 881 LeuLleLeuTyrSerAenAenHisThrGlnProGluHisGlnTrpThrLeuLeuThrThr 900
Db 2641 CTAAATCTCTACAGCAACCAACACACCGCGCCGCAACACCAAGTGGACATCTCCACCA 2700
Qy 901 GluGlyAenLlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePhe 920
Db 2701 GAGGGAAACATCTTCAGTTCAGAGGTCCATGGCCCTAGAGAGTGACACTCGGTATTTCTC 2760
Qy 921 LysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValile 940

Db 2761 AAGATGGAGCCCGCACAGAGGTGGGCGCTTGGGCGCTTTTCCGCTTGCAGGATGTGATT 2820
Qy 941 ThrLeuGlnGlnThrPheSerAspSerLeuAspValHisAlaValThrGlyLleileVal 960
Db 2821 ACTCTGCAAGAGACATTTCTCAGACTCTCTGGATGTGCACGCGCTCGTCGGGCATCATCTGT 2880
Qy 961 GlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGlnSer 980
Db 2881 GGTGTCTGGCTGGGCGCTTCTCTGCTCTCTGGCTTGCATGTGTCTGTGCGCTACGACAAAGC 2940
Qy 981 SerHisArgGluAlaLeuProGlyLeuSerSerSerGlyThrProGlyAenProAlaLeu 1000
Db 2941 TCCACAGGAGAGCGCTTCCCGATTTGTCTCTCAGGCAACCCAGGAAACCCAGCGCTC 3000
Qy 1001 TyrThrArgAlaArgLeuGlyProProSerValProAlaAlaHisGluLeuGluSerLeu 1020
Db 3001 TACACAAGAGCTCGCTTGGGCTCTCCAGTGTCTCTGCTGCCATGATGTTGGAGTCCCTC 3060
Qy 1021 ValHisProArgProGlnAspTrpSerProProSerAspValGluAspLysAlaGlu 1040
Db 3061 GTGCATCTCTGCTCCCGAGGATTTGGTCCCAACCCCTCAGATGTGGAAGAACAAGGCTGAA 3120
Qy 1041 ValHisSerLeuMetGlyGlySerValSerAspCysArgGlyHisSerLysArgLysile 1060
Db 3121 GTACACAGCTTATGGTGGCAGTGTCTCAGATTGCCGGGCGCACTCCAAGAGAAAGATC 3180
Qy 1061 SerTrpAlaGlnAlaGlyGlyProAenTrpAlaGlySerTrpAlaGlyCysGluLeuPro 1080
Db 3181 TCCTGGGCTCAGGCAGCGGAGCAAACTGGGCGAGGCTCTGGGCGAGCTGTGAGCTGCC 3240
Qy 1081 GlnGlySerGlyProArgProAlaLeuThrArgAlaLeuLeuProProAlaGlyThrGly 1100
Db 3241 CAGGATAGTGTCCAAAGCGCGCTCTGACCCCTGCTCTGCTGCTCAGCGGGAACCGGG 3300
Qy 1101 GlnThrLeuLeuLeuGlnAlaLeuValTyrAspGlyLleLysSerAenGlyArgLysile 1120
Db 3301 CAGACACTGCTGTCTGCAAGCGCTCTGGTGTATGACGCGATAAGCAACCGGGAAGAAG 3360
Qy 1121 ProSerProAlaCysArgAenGlnValGluAlaGluValLleValHisSerAspPheGly 1140
Db 3361 CCGTCCCGCCGCTCGAGGAATCAGTGGAGTGGAGTCAATTGTCTCACTCCGACTTCGGT 3420
Qy 1141 AlaSerLysGlyCysProAspLeuHisGlnAspLeuGluProGluGluProLeuThr 1160
Db 3421 GCATCCAAAGATGTCTCTGACCTCCACTCCAAGACCTGGAGCCAGAGGAACCACTGACT 3480
Qy 1161 AlaGluThrLeuProSerThrSerGlyAlaValAspLeuSerGlnGlyAlaAspTrpLeu 1180
Db 3481 GCAGAGACTCTGCTTCCAGCTCTGGAGCTGTGGATCTGTCTCAAGGAGCAGACTGGCTG 3540
Qy 1181 GlyArgGluLeuGlyCysGlnProThrThrSerGlyProGluArgLeuThrCysLeu 1200
Db 3541 GGCAGGAGCTGGGAGGTGTCACCAACCAACGAGTGGGCGCAGAGAGGCTCACTGTCTG 3600
Qy 1201 ProGluAlaAlaSerAlaSerCysSerCysAspLeuGlnProSerThrAlaLleGlu 1220
Db 3601 CCAGAGCAGCAGTGGCTCTCTGCTCTGCTCTCAGACCTCCAAGCCAGCACTGTCTATAGAG 3660
Qy 1221 GluAlaProGlyLysSerCysGlnProLysAlaLeuCysProLeuThrValSerProSer 1240
Db 3661 GAGGCGCTTGGGAAAGCTCCAGCCCAAGCGCTGTGTCTCTTAAAGTCAAGTCAAGCCCAAGC 3720
Qy 1241 LeuProArgAlaProValSerSerAlaGlnValPro 1252
Db 3721 CTTCCAGGCGCTCTGTCTCTCTCTGCTCAGGTCCCC 3756

RESULT 2

AX191259
LOCUS AX191259
DEFINITION Sequence 1 from Patent W00149714.
ACCESSION AX191259
VERSION AX191259.1 GI:15209510
KEYWORDS

AX191259 6176 bp DNA linear PAT 15-AUG-2001

SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE Salbaum, J. M.
AUTHORS Nope polypeptides, encoding nucleic acids and methods of use
TITLE Patent: WO 0149714-A1 12-JUL-2001;
JOURNAL Neurosciences Research Foundation Inc. (US)
FEATURES Location/Qualifiers
source 1..6176
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
1..3759
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC51120.1"
/db_xref="GI:15209511"
/translation="MARADTGRGLLVLTFCLLSARGELPLPORTTIVKLSCDRGLPQVI
LGPEQAVLDCTLGATAAGPPRTVWSKGDVTLHENLHLLPNSLSMLSSPLEQDS
DEEALRIWKVTEGTSYCLAHSPLGVVSAQVAVKLTUDEFSLPESQIVBENGAR
FECHTQAPAIITWKGQVTPVEPRLLITLPMKLLQILDVQDSAGSYRCVATNSAR
QFSSQASLTVALRGLEATRGQDVVIAAPENTVVSQNVMBCVASADPTPVSW
VRQDKPISTDVILGRTNLLTASQPRHSVVCRANKPLTRDPATAAELRLVAAP
AISOAPEALSRTASTARFVCRASGEPRPALHLDGIPLRPNRGVKGQGGSLVIT
QIGLDAGYQCVAENSAGTACAAFLAVVREGLPSPATRVATPLSSSVLVAMER
PELHSEQIIGFSLHQRKGVNDVEYQFVANNDDITELQVRDLSPNDTDFYVAYSQI
GASRTSPALVHTDDVPAAPQLTSSPNSDIRVAMLPFSSISNGQVLYKIBYG
LQKEDQVFTEVPGNETQTLNSLPQNKVYRISAGTGAGYGVSPQMQRHTPGVHN
QSHVPPAELKVRAMKESLVSMQPPHPTQISGKLYMGVGTGEEBADGDPGGR
GQADWDGVRLLKVKQYELTQVPRVEVKLVAFNKHEDGYAAVWGKTEKAPT
DLPQGRPELPHNVHAESENSSTSIWLRNKKDPFTVKLVNTVTRFPGWGLNASLVT
YITSGEDLILGKFTKTFEAVQSGVMDGPPGVSVVERSTLPDRPSTPPSDLRLS
PLTPSVRLHMCPTPENGEIVYLLYSNNHTQPEHOWTLTTEGNTFSARVHGLS
DTRYFFKMGARTGEPFSLQDVTITLOETFSDSLDDHVAVTGIIIVGVCGLLCLAC
MCAQLQSSHREALPGLSSGTPGNALYTRALRGPSPVPAHAHELSLVHPRPQDWS
PSDVEDKAEVSLMGSSVSDCRGHSKRKISMAQAGPNWAGSWAGCELPQSGRPPA
LYRALLPAGTQGLLQALVYDGKSNRGRKPSACRNOQVRAEVIVHSDFGASKGCP
DLHLQDEPPEPITASTLPSGAVDLSQADWLGRGLGCGQPTTSGPRLTCLPEAA
SASCSDLPQSPSTAIEAPFKSCQPKALCPLTVSPSLPRAPVSSAQVP"
BASE COUNT 1316 a 1756 c 1634 g 1470 t
ORIGIN
Alignment Scores:
Pred. No.: 3,04e-260 Length: 6176
Score: 6604.00 Matches: 1252
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-754-997A-2 (1-1252) x AX191259 (1-6176)
QY 1 MetAlaArgAlaAphThrGlyArgGlyLeuLeuValLeuThrPheCysLeuLeuSerAla 20
DB 1 ATGCGCGGGCGGACAGCGCGCGGCTCCTGGTGTGACCTTCTGCGCTGTGCTCGCG 60
QY 21 ArgGlyGluLeuProLeuProGlnGluThrThrValIysLeuSerCysAspGluGlyPro 40
DB 61 CGCGGGGAGGTGCATTCGCCAGGAGACAACTGTCAAGCTGAGCTGTGATGAGGACCC 120
QY 41 LeuGlnValIleLeuGlyProGluGlnAlaValValLeuAspCysThrLeuGlyAlaThr 60
DB 121 CTGCAAGTGTATCTCGGCCCTTGACAGAGCTGTGTGTGCTGACCTGCTTTGGGGGCTACA 180
QY 61 AlaAlaGlyProProThrArgValThrTrpSerIysAspGlyVspThrValLeuGluHis 80
DB 181 GCTGCTGGGGCTCCGACACAGGGGTGACATGGAGCAAGATGGAGCACTGTACTAGACAT 240
QY 81 GluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGluGlnGlu 100
DB 241 GAGAACTGCACTGTCTACCCANTGGCTCCCTGTGCTGCTCTCCCTACCCCTAGACAGAA 300

101 ApsSerAspAspGluGluAlaLeuArgIleTrpIysValThrGluGlySerTrpSerCys 120
301 GACAGCGATGATGAGAGAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTCTGT 360
121 LeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValValIysLeuAlaThr 140
361 CTGGCCCAACAGCCCGCTAGGAGTGTGGCCAGCCAGGTGTCTGTGTCTCAAGCTTCCACA 420
141 LeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThrAlaArg 160
421 CTGGAAGACTTCTCTCTGCACCCGAGTCCAGATGTGGAGAGAGAACGGGACACGACGC 480
161 PheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluIysAspGlnVal 180
481 TTTGTAATGCCACACCAAGGGCTTCCAGCCCCCATCATCTTCTGGGAAAGGACCCAGGTG 540
181 ThrValProGluGluProArgLeuIleThrLeuProIysTrpLeuLeuGlnIleLeuAsp 200
541 ACCGTGCTCGAGGAGCCCGGCTCATCACTCTTCCCAAGTGTCTCTCCAGATCCTAGAT 600
201 ValGlnAspSerAspAlaGlySerTrpArgCysValAlaThrAsnSerAlaArgGlnArg 220
601 GTCCAGGACAGTGTATGCGAGGTCTTACCGCTGCGTGGCCACCAATTTCAGCCCGCCCAACGA 660
221 PheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGly 240
661 TTCAGCCAGGAGCCCTCGCTCATCTGTGGCCCTCAGAGGGTCTTTTGAGGCTTACAGGGGG 720
241 GlnAspValValIleValAlaAlaProGluAsnThrThrValValSerGlyGlnAsnVal 260
721 CAGGATGTGGTCAATTTGTGGCAGCCCAAGAACACCAAGGTAGTGTCTGGACAGAAATGTA 780
261 ValMetGluCysValAlaSerAlaAspProThrProPheValSerTrpValArgGlnAsp 280
781 GTGATGGAGTGGCGTGGCTCTGTCTACCCACCCCTCTTGTGTCTCTGGGTCCGACAGGAT 840
281 GlyLysProIleSerThrAspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSer 300
841 GGAAGACCTATCTCCAGATGTCTGTCTGGGCGGACCACTTACTCTATCGCCAGC 900
301 AlaGlnProArgHisSerGlyValTrpValCysArgAlaAsnLysProLeuThrArgAsp 320
901 GCGCAGCTCGGCACTCTGGAGTCTATGTCTGCGAGCCCAAGCCCTCAGCGGTGAC 960
321 PheAlaThrAlaAlaAlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaPro 340
961 TTTCCGCACTGCGCGCTGTGAGCTCCGAGTGTCTGCTGCCCGCCAGCCATCTCGCAGGCACC 1020
341 GluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGlu 360
1021 GAGGCGCTCTCGGAGACCGCGGCGCAGCACCGCGGCTTGTGTGCGGCGGTCCGCGGAG 1080
361 ProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgProGlnGlyArgVal 380
1081 CCACGCGCGCGCTGCACCTGGCTGCAGCAGCGGATCCCGTTCGACCCCAATGGCGCGCTC 1140
381 LysValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGly 400
1141 AAGGTGCGAGCGGTGGCGGAGCTTGTGTCTCATCTCAGATCGGCTGCGGCTGCGGAGCGTGGC 1200
401 TyrTrpGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaAlaProLeuAla 420
1201 TACTACAGTGTGAGAGAAACGCGGGAACCTGCTGTGCGGCTGCGGCTGCGGCTGCGG 1260
421 ValValValArgGluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSer 440
1261 GTAGTGTGCGGAGGGGCTGCGAGCGCCCGACTCGGCTCAGCCACGCGCGCTGAGC 1320
441 SerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGluGlnIleLeuGly 460
1321 AGCTCTCTGTGTGTGGCTGGGAGCGGCTGAGTTGACAGCGAGCAAAATCAATGGC 1380

QY 461 PheSerLeuHisTyrGlnLysAlaArgGlyValAspAsnValGluTyrGlnPheAlaVal 480
DB 1381 TTCTCTCTTCACTACCAAAAGCAAGGGAGTGGCAATGTGGAGTACAGATTGACGTA 1440
QY 481 AsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPhe 500
DB 1441 AACAAATGACACACACAGAGCTGCGAGTTTGGGACCTGGAAACCCACACCGGATTTAGATTTC 1500
QY 501 TyrValValAlaTyrSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHis 520
DB 1501 TACGTGGTGGGCTACTCCAGCTGGGGGCCAGCCGAACTCCAGGCCAGCTGGTGCAAT 1560
QY 521 ThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSer 540
DB 1561 ACACGTGGACAGATGTCCTCCAGCGCAGACCCAGCTTACTTGTTCAGGCCCAACCCCTCG 1620
QY 541 AspIleArgValAlaTyrProLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLys 560
DB 1621 GACATCAGGTGGCATGTGCTGCCCTGCCCTCCAGCTGAGCAATGGACAGTGTCTGAG 1680
QY 561 TyrIleIleGluTyrGlyLeuGlyLysGluAspGlnValPheSerThrGluValProGly 580
DB 1681 TACAAAGATAGAGTACGGTTTGGGGAAGGAGATCAGGTTTCTCCACCGAGGTGCTGGA 1740
QY 581 AsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgIle 600
DB 1741 AATGAGACACAACTTACGTATAACTCACTCAGGCCAAACAAAGTGTACCGAGTCCGGATT 1800
QY 601 SerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTyrMetGlnHisArgThrPro 620
DB 1801 TCAGCTGGCAGCTGGCGCTGAGTATGGAGTCCCTCTCAGTGGATGACAGACAGACCT 1860
QY 621 GlyValHisAsnGlnSerHisValProPheAlaProAlaGluLeuLysValArgAlaLys 640
DB 1861 GGTGTGCACACACAGAGCAGTGTCTCTTGGCCCTGCAGAAATTGAAGTGAAGGCAAG 1920
QY 641 MetGluSerLeuValSerTyrGlnProProHisProThrGlnIleSerGlyTyr 660
DB 1921 ATGGAGTCCCTGGTGGTGTATGGCAGCGCGCCCTCACCCACCCAGATCTCTGGATAC 1980
QY 661 LysLeuTyrTyrGlyValGlyThrGluGluAlaAspGlyAspArgProProGly 680
DB 1981 AAACCTTACTGGGAGAGGTGGAAACAGAGGAGGAGGAGATGGTGAACCGCCCCCAGGG 2040
QY 681 GlyArgGlyAspGlnAlaTyrAspValGlyProValArgLeuLysLysValLysGln 700
DB 2041 GGTGTGGAGATCAGCTTGGAGCTCGGGCCCTGGCGCTGAGAAAGAAAGTGAAGCAG 2100
QY 701 TyrGluLeuThrGlnLeuValProGlyArgProTyrGluValLysLeuValAlaPheAsn 720
DB 2101 TATGAACCTGACCCAGTTAGTCCCTGGCAGGCGGTACGAGGTGAGCTCGTAGCTTTCAAC 2160
QY 721 LysHisGluAspGlyTyrAlaAlaValTyrLysGlyLysThrGluLysAlaProThrPro 740
DB 2161 AAAACAGAGGACGGCTACGCTGCTGTGTGGAAGGCAAGACGAGGAGCGCCCAAGCCA 2220
QY 741 AspLeuProfileGlnArgGlyProProLeuProProAlaHisValHisAlaGluSerAsn 760
DB 2221 GACCTGCCCTATCCAGAGGGGGCCACCGCTGCCCTCTGCCCATGTCCAGCAGAGTCAAC 2280
QY 761 SerSerThrSerIleTyrLeuArgTyrLysLysProAspPheThrThrValLysIleVal 780
DB 2281 AGCTCCACTTCCATTTGCTTCCGTGGGAAGACAGACTTACCACTGTCAAGATTGTC 2340
QY 781 AsnTyrThrValArgPheGlyProTyrGlyLeuArgAsnAlaSerLeuValThrTyrTyr 800
DB 2341 AACTACACTGTACGCTTCCGGCCCTCGGGGCTCAGGAATGCTTCCCTGGTCACTACTAT 2400
QY 801 ThrSerSerGlyGluAspIleLeuIleGlyLysLysProPheThrLysTyrGluPhe 820
DB 2401 ACCAGCTCTGGAGAGACATTTCTCATTTGGCGGCCCTGAAACCACTTTACCAAGTACAGTTT 2460
QY 821 AlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSer 840

DB 2461 GCGGTACAGTCCCAACGAGTGGATATGGATGGGCCCTTTGGCTCCGTAGAACGCTCC 2520
QY 841 ThrLeuProAspArgProSerThrProSerAspLeuArgLeuSerProLeuThrPro 860
DB 2521 ACCCTGCTGACCGGCTTCAACACCTCTTCTGACCTGGCCCTGAGGCCCTGACACCA 2580
QY 861 SerThrValArgLeuHisTyrProProProThrGluProAsnGlyGluIleValGluTyr 880
DB 2581 TCCACCGTTCGGTTTACATCTGTGTCCCCCAACGAGGCCAATGGTGGAGATTGTGGAGTAT 2640
QY 881 LeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTyrThrLeuLeuThrThr 900
DB 2641 CTAAATCTCTACAGCAACACACACCGAGCCCAACACCAAGTGGACACTCTCTCACCACA 2700
QY 901 GluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePhe 920
DB 2701 GAGGGAAACATCTTCAGTGCAGAGTCCATGGCCCTAGAGAGTGACACTCGGTATTCTTC 2760
QY 921 LysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIle 940
DB 2761 AAGATGGAGCGCCACACAGAGTGGGCCCTGGGCCCTTTTCCGCTTGCAGGATGTGATT 2820
QY 941 ThrLeuGlnGluThrPheSerAspSerIleuAspValHisAlaValThrGlyIleIleVal 960
DB 2821 ACTCTGCAAGAGACATTTCTCAGACTCTCTTGGATGTGACGCGCTCAGCGGCATCATCGT 2880
QY 961 GlyValCysLeuGlyLeuLeuCysLeuAlaCysMetCysAlaGlyLeuArgGlnSer 980
DB 2881 GGTGTCTGCCCTGGGCCCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940
QY 981 SerHisArgGluAlaLeuProGlyLeuSerSerSerGlyThrProGlyAsnProAlaLeu 1000
DB 2941 TCCACAGGGAAGCCCTTCCCGGATTGTCTCTCAGGCACCCCGAGGAACCCAGCGCTC 3000
QY 1001 TyrThrArgAlaArgLeuGlyProProSerValProAlaAlaHisGluLeuGluSerLeu 1020
DB 3001 TACACAAGAGCTCGGCTTGGGCCCTCCAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 3060
QY 1021 ValHisProArgProGlnAspTyrProProProSerProSerProSerProSerProSerPro 1040
DB 3061 GTGCATCTCTGCTCCAGAGATTGGTCCCAACCTCAGATGTGAAGAAGACAGGCTGAA 3120
QY 1041 ValHisSerLeuMetGlyGlySerValSerAspCysArgGlyHisSerLysArgLysIle 1060
DB 3121 GTACACAGCCTTATGGTGGCAGTGTTCAGATTGCCGGGCCCTCCACAGAGAAGATC 3180
QY 1061 SerTrpAlaGlnAlaGlyProAsnTrpAlaGlySerTrpAlaGlyCysGluLeuPro 1080
DB 3181 TCCTGGGCTCAGGCGGGGACCAAACTGGGCGAGGCTCTGGGCGAGGCTGTGAGCTGCC 3240
QY 1081 GlnGlySerGlyProArgProAlaLeuThrArgAlaLeuLeuProProAlaGlyThrGly 1100
DB 3241 CAGGGTATGTGTCCAAAGCGCGCTCTGACCCGCTCTGCTGCTCAGCGGGAACCGGG 3300
QY 1101 GlnThrLeuLeuLeuGlnAlaLeuValTyrAspGlyIleLysSerAsnGlyArgLysLys 1120
DB 3301 CAGACACTGTCTGTGCAAGCCCTGTGTATGACGGCATTAAGAGCAACCGGAGAAAGAG 3360
QY 1121 ProSerProAlaCysArgAsnGlnValGluAlaGluValIleValHisSerAspPheGly 1140
DB 3361 CCGTCCCAGGCTCAGGAATCAGGTGGAGCTGAGGTCAATTGTCCACTCCGACTTCGCT 3420
QY 1141 AlaSerLysGlyCysProAspLeuHisLeuGlnAspLeuGluProGluGluProLeuThr 1160
DB 3421 GCATCCAAAGGATGTCTCTGACTCCACTCCAGACCTCGAGCGCAGGAGGAAACCACTGACT 3480
QY 1161 AlaGlnThrLeuProSerThrSerGlyValAlaAspLeuSerGlnGlyValAspTrpLeu 1180
DB 3481 GCAGAGACTCTGCCCTCCAGCTCTGGAGCTGTGTATCTGTCTCAAGGAGCAGATGCGCTG 3540
QY 1181 GlyArgGluLeuGlyGlyCysGlnProThrThrSerGlyProGluArgLeuThrCysLeu 1200

Qy	261	ValMetGluCysValAlaSerAlaAspProThrProPheValSerTrpValArgGlnAsp	280
Db	781	GTGATGAGTGCCTGGCGCTCTGCTGACCCACCCCTTTGTGTCTGGTCCGACAGGAT	840
Qy	281	GlyIysProIleSerThrAspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSer	300
Db	841	GGAAAGCCTATCTCCACGGATGTCNTCGTTCTGGCCGGACCAATCTACTCATCGCCAGC	900
Qy	301	AlaGlnProArgHisSerGlyValTyrValCysArgAlaAsnLysProLeuThrArgAsp	320
Db	901	CGCGAGCCTCGGCNCTCTGGAGTCTATGTCTGCGAGCAACAGGCCCTCACCGCTGAC	960
Qy	321	PheAlaThrAlaAlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaPro	340
Db	961	TTGCCACTCGCGCTGCTGAGCTCCGAGTCTGCTGCTGCCCGAGCATCTCGCAGGCACCC	1020
Qy	341	GluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGlu	360
Db	1021	GAGCGCTCTCGCGGACCGCGGCGCAGCACCGCGCGCTTCGTGTGCCGGCGGTCCGGGAG	1080
Qy	361	ProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgVal	380
Db	1081	CCACGGCCCGCGCTGCACTGGCTGCACGACGGGATCCCGTTGGACCCCATGGCGGCTC	1140
Qy	381	LysValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGly	400
Db	1141	AAGGTGACGGCGGTGGCGCAGCTTGCTCATCACTCAGATCGGCTGCAGGACGCTGGC	1200
Qy	401	TyrTrpGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaAlaProLeuAla	420
Db	1201	TACTACAGTGCCTAGCAGAAAACAGCGCGGAACTGCTCTGTCGCGTGCGCCCTCGCG	1260
Qy	421	ValValValArgGluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSer	440
Db	1261	GTAGTGTGTGCGAGGGGCTGCCAGCGCCCGGACTCGGGTCAAGCCACCGCGCTGAGC	1320
Qy	441	SerSerSerValLeuValAlaIleTrpGluArgProGluLeuHisSerGluGlnIleIleGly	460
Db	1321	AGCTCCTCTGTGCTGTGGCTGGAGCGCGCTGAGTTGCACAGCAGCAAAATCATTTGGC	1380
Qy	461	PheSerLeuHisTrpGlnLysAlaArgGlyValAspAsnValGluTyrGlnPheAlaVal	480
Db	1381	TTCTCTCTTCACTACCAAAAGGCAAGGGGAGTGGCAATGTGGAGTACCAAGTTTGAGTA	1440
Qy	481	AsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPhe	500
Db	1441	AACNATGACACCAACAGAGCTGCAGGTTCCGGACTTGGNACCCAAACGGAATTATGAGTTC	1500
Qy	501	TyrValValAlaTyrSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHis	520
Db	1501	TAGTGTGTGGCTACTCCAGCTGGGGGCGACCGCAACCTCCAGCCAGCCCTGGTGCAT	1560
Qy	521	ThrLeuAspAspValProSerAlaAlaProGlnLeuThrIleuSerSerProAsnProSer	540
Db	1561	ACACTGACAGATGTCCCACGCGCAGCACCCCGACTTACCTTGTTCAGGCCCAACCCCTCG	1620
Qy	541	AspIleArgValAlaIleTrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLys	560
Db	1621	GACNATAGGGTGGCATGTGCTGCCCTCCAGCTTGAGCAATGACAGGTGCTGGAAG	1680
Qy	561	TyrLysIleGluTyrGlyLeuGlyLysGluAspGlnValPheSerThrGluValProGly	580
Db	1681	TACAAGATAGAGTACGGTTTGGGGNAGGAAGATCAGGTTTTCTCCACCGAGGTGCCTGGA	1740
Qy	581	AsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgIle	600
Db	1741	AATGAGACACACTTACGTTTAAACTCACTTCAGCCAAACAAAGTGTACCGAGTCCCGATT	1800
Qy	601	SerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrPro	620
Db	1801	TCAGCTGGCACTGGCGCTATGAGTCCCTTCTCAGTGGATGTCAGCACAGACACCT	1860
Qy	621	GlyValHisAsnGlnSerHisValProPheAlaProAlaGluLeuLysValArgAlaLys	640

QY 61 AlaAlaGlyProProThrArgValThrTrpSerLysAspGlyAspThrValLeuGlnHis 80
DB 230 GCTGTGGGCTCCACAGAGGTGACATGGAGCAAGGATGGAGACACTGTATAGAGCAT 289
QY 81 GluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGlnGln 100
DB 290 GAGAACCTTGCACTGCTACCCAAATGGCTCCCTGTGGCTGTCTCTACCCCTTAGAGCAAGAA 349
QY 101 AspSerAspGluGluAlaLeuArgIleTrpLysValThrGluGlySerTrpSerCys 120
DB 350 GACAGCGATGATGAGGAAGCTCTTAGGATCTGGAAAGTCACTAGAGGCGAGCTATTCTCTGT 409
QY 121 LeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValValLysLeuAlaThr 140
DB 410 CTGGGCCACAGCCCTTAGGAGTGGTGGCCAGCCAGGTGTGCTGTGGTCAAGCTTGCCACA 469
QY 141 LeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThrAlaArg 160
DB 470 CTCGAAGACTTCTCTGCAACCCGAGTCCCAAGATTGGAGGAGAACGGGACAGCAGC 529
QY 161 PheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAspGlnVal 180
DB 530 TTGGAATGCCACACCAAGGCTCTCCAGCCCCCATATTACTTGGAAAGGACCAAGTG 589
QY 181 ThrValProGluGluProArgLeuIleThrLeuProLysTrpLeuGlnIleLeuAsp 200
DB 590 ACCGTGCTGAGGAGTCCGGCTCATCTCTTCCCAATGGCGTCTCCAGATCCTAGAT 649
QY 201 ValGlnAspSerAspAlaGlySerTrpArgCysValAlaThrAsnSerAlaArgGlnArg 220
DB 650 GTCCAGGACAGTATGTCAGGCTCTTACCCTGCGTGGCCACCAATTCAGCCCGCCCAACGA 709
QY 221 PheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGly 240
DB 710 TTCAGCCAGAGGCTCTGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTACCAAGGGG 769
QY 241 GlnAspValValIleValAlaAlaProGluAsnThrThrValSerGlyGlnAsnVal 260
DB 770 CAGGATGTGGTCAATTTGTGGCAGCCCGCAGAGAACACACCGTAGTGTCTGGACAGAGTGA 829
QY 261 ValMetGluCysValAlaSerAlaAspProThrProPheValSerTrpValArgGlnAsp 280
DB 830 GTGATGGAGTGGTGGCTCTGCTGACCCCACTTGTGTCTGGTCCCGACAGGAT 889
QY 281 GlyLysProIleSerThrAspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSer 300
DB 890 GGAAAGCCTATCTCCAGGATGTCATCGTTCTGGGCCCGACCAATCTACTCATGCCACG 949
QY 301 AlaGlnProArgHisSerGlyValTrpValCysArgAlaAsnLysProLeuThrArgAsp 320
DB 950 GCGCAGCCTCGGCACTCTGGAGTCTATGTCTGCCGAGCCAAAGCCCGCAGCGCTGAT 1009
QY 321 PheAlaThrAlaAlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaPro 340
DB 1010 TTCGCACTCGCGGCTGTGAGTCCGAGTGTGCTGTGCCCCAGCCCATCTCATGCCACG 1069
QY 341 GluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGlu 360
DB 1070 GAGGCGCTCTCGGACCGGGCCAGCACCGCGCTTCTGGTGGCGGGCTGCGGGGAG 1129
QY 361 ProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgVal 380
DB 1130 CCACGGCCCGCTGCACTGCTGACCGAGGATCCCGTTGGCAGCCCAATGGGCGCTC 1189
QY 381 LysValGlnGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGly 400
DB 1190 AAGGTGACGGGCGGTGGGGCGAGCTTGTGTCATCACTCAGATCGGCTTTCAGGACGCTGGC 1249
QY 401 TyrTrpGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaProLeuAla 420
DB 1250 TACTACAGTGTGATAGCAAAACAGCGGGAACTGCTGTGGCGTGGCCCTGGCG 1309

QY 421 ValValValArgGluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSer 440
DB 1310 GTAGTGTGGCGAGGGGCTGCCAGCGCCCCGACTCGGGTCAACAGCCACCCCGCTGAGC 1369
QY 441 SerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGly 460
DB 1370 AGCTCTCTGTGCTGTGGCTGGAGCGGCTGAGTTGGACACGAGCAAAATCATTTGGC 1429
QY 461 PheSerLeuHisTrpGlnLysAlaArgGlyValAspAsnValGluTrpGlnPheAlaVal 480
DB 1430 TTCTCTCTTCACTACCAAAAGGCAAGGGAGTGGACAAATGTGGAGTACCACTTTTGCAGTA 1489
QY 481 AsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTrpGluPhe 500
DB 1490 AACAAATGACACACAGAGCTGAGGTTGGGACCTTGGAAACCAACACGAGTATTAGTTTC 1549
QY 501 TyrValValAlaTrpSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHis 520
DB 1550 TACGTGTGGCTTACTCTCCAGCTGGGGCCAGCCGAACCTCCAGCCAGCCCTGGTGAT 1609
QY 521 ThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSer 540
DB 1610 ACCTGGAGCATGTCCCAGCGCAGCACCCAGCTTACCTTGTCCAGCCCCCAACCCCTCG 1669
QY 541 AspIleArgValAlaTrpLeuProLeuProSerSerSerLeuSerAsnGlyGlnValLeuLys 560
DB 1670 GACATCAGGGTGGATGGCTGCCCTCTCCAGCTGGAGCAATGAGCAATGAGCAGTCTGAAG 1729
QY 561 TyrLysIleGluTrpGlyLeuGlyLysGluAspGlnValPheSerThrGluValProGly 580
DB 1730 TACAAAGATAGAGTACGGTTTGGGGAAGGAAGATCAGGTTTTCTCCAGGAGTGGCTGGA 1789
QY 581 AsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTrpArgValArgIle 600
DB 1790 AATGAGACACAACTTACGTTAACTCACTTCAGCCAAACAAAGTGTACCGAGTCCGAGTT 1849
QY 601 SerAlaGlyThrGlyAlaGlyTrpGlyValProSerGlnTrpMetGlnHisArgThrPro 620
DB 1850 TCAGCTGGCACTGGCGCTGGCTATGGAGTCCCTTCTCAGTGGATGTCAGCACAGCACCT 1909
QY 621 GlyValHisGlnGlnSerHisValProPheAlaProAlaGluLeuLysValArgAlaLys 640
DB 1910 GGTGTGCACAAACAGAGCCATGTTCCTTTGCCCTCTCAGAAATTTGAAGGTGAGGGCAAG 1969
QY 641 MetGluSerLeuValValSerTrpGlnProProHisProThrGlnIleSerGlyTrp 660
DB 1970 ATGGAGTCCCTGGTGGTGTCAATGGCAGCGCCCTCTACCCCAACCCAGATCTCTGGATAC 2029
QY 661 LysLeuTrpTrpGlyGluValGlyThrGluGluAlaAspGlyAspArgProProGly 680
DB 2030 AAACCTCTACGAGAGAGTGGGAAACAGAGAGGAGGAGAGTGGTGCAGCGCCCGCCAGGG 2089
QY 681 GlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLysLysValLysGln 700
DB 2090 GGTGTGGAGATCAAGCTTGGGACGTGGGCGCCGTGGCGCTGGAAGAAAGATGAAGCAG 2149
QY 701 TyrGluLeuThrGlnLeuValProGlyArgProTrpGluValLysLeuValAlaPheAsn 720
DB 2150 TATGAACTGACCCAGTTAGTTCCTTGGCAGGCTGTACGAGGTGAAGCTCTGTAGCTTTCAAC 2209
QY 721 LysHisGluAspGlyTrpAlaAlaValTrpLysGlyLysThrGluLysAlaProThrPro 740
DB 2210 AAACACGAGGACGGCTACGCTGCTGTGTGGAGGCAAGACGAGAGAGGCGCCACCGCCA 2269
QY 741 AspLeuProIleGlnArgGlyProProLeuProProAlaHisValHisAlaGluSerAsn 760
DB 2270 GACCTGCTATTCAGAGGGGCGCCACCGCTGCTCTCCCAATGTCCACGAGAGTCAAAAC 2329
QY 761 SerSerThrSerIleTrpLeuArgTrpLysLysProAspPheThrThrValLysIleVal 780
DB 2330 AGCTCCACTTCCATTTGGCTTGGTGGAGAGGACGACACTTTACCACTGTCAAGATTGTC 2389
QY 781 AsnTrpThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTrpTyr 800

VROGKPISTDVIVLGTETNLASQPHSGVYVCRANKPRTRDPATATAAELRLVLAAP
 AISQPEALSTRASTARFVCRASGPPALHMLHDGIPLRPNGRKVRQGGGSLIVIT
 QIGLQDAGYQCVASNSAGTACAAAPVVRREGLPFSAPTRVATPISSSSVLVAMER
 PELSHQYIGLSHYKQAGVDNVEYQFVANNDDTTELQVRLDRLEPNTDFFYFVAYISOL
 GASRTSSPALVHTLDDVPSAAPQLTSSPNPSDIRVAMLPSPSSLSNGOMLKYKIBYG
 LKGEEDOVSTEVPGNETOLTSLQPNKVRVRIISACTGAGYGVPSOMOHRTPGVH
 NOSHVPPAPAEKLVRAKWEISLVSNQPPHPPTQISGYKLYREVGTEBERADGRPPGG
 RGDQADVGVRLKVKVQYBELTQVLPGLRYEVLKLVAFNKHEDGTAAVWKGKTKKAPT
 PDLPIQRPPLPPAHVHAESNSSTISILRWKKPDFTTVKIVNIVTRFPGMGRNASIV
 TYTSSGSDILIGLKKPTKYEFVQSHGVDMGDPFGSVVERSTLPPRSPFPDGLRL
 SPLTSPSTVRLHWCPTBPNGEIVELYLLYSNNHTQPEHOWTLLATTEGNI FSAHVHGLR
 SDTRYFFKMGARTVEGPPFRLODVLITLOKTSFDSLNDVHAVTGII VGCVLGLLCLLA
 CMKAGLRSSHREALPGLSSSGTFCNFPALYTRALRGLPPSPAHELSLHVPRQDWS
 PPSDVEDKASHVSLMGSSVSDCRGSKKLSWAQAGFPNWSAGWAGCELPQGGSPRP
 ALTRALLPAGTGTLLQALVYDAIKNSGRKPSFACRNQVREAVIIVSHDFPASKGC
 PDLHLQDLEPEPTTAFTLPSTGAVDLSQGDWLGRELGCQPTTSGPRLTCLPEA
 ASACSCSDLPQSTAIERAPGKSCQPKALCPLTVSPSLPRAPVSSAQPV"

BASE COUNT 1321 a 1772 c 1653 g 1476 t

ORIGIN

Alignment Scores:

Pred. No.: 7,03e-257 Length: 6222
 Score: 6521.50 Matches: 1241
 Percent Similarity: 99.36% Conservative: 4
 Best Local Similarity: 99.04% Mismatches: 7
 Query Match: 98.75% Indels: 1
 DB: 10 Gaps: 1

US-09-754-997A-2 (1-1252) x AB052621 (1-6222)

Qy	1	MetAlaArgAlaAspThrGlyValLeuValLeuThrPheCysLeuLeuSerAla	20
Db	50	ATGGCGCGGGGACACACGGCGCGGCTCTGGTGTGACCTTCCTGCTGCTGCGCG	109
Qy	21	ArgGlyGluLeuProLeuProGlnGluThrValLeuLeuSerCysAspGluClyPro	40
Db	110	CGCGGGAGCTGCGCATTCGCCGAGGACAACTGTCAAGCTGAGCTGTGATGAGGACCC	169
Qy	41	LeuGlnValIleLeuGlyProGluGlnAlaValValLeuLeuAspCysThrLeuGlyAlaThr	60
Db	170	CTGCAAGTGATCTTGGGCGCTGAGCAGGCTGTGGTGTGACTTGGGCGCTACA	229
Qy	61	AlaAlaGlyProProThrArgValThrTrpSerLysAspGlyAspThrValLeuGluHis	80
Db	230	GCTGTGGGCGCTCCGACAGGCTGACATGAGCAAGGATGGAGACACTGTACTAGAGCAT	289
Qy	81	GluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGluGlnGlu	100
Db	290	GAGAACCTGACCTCTACCAATGGCTCCCTGTGTGGCTGTCTTACCCCTTAGAGCAAGA	349
Qy	101	AspSerAspAspGluGluAlaLeuArgIleTrpLysValThrGluGlySerTyrSerCys	120
Db	350	GACAGCGATGATGAGAGAGCTTTAGGATCTGGAAGGTCACTAGGCGGAGCTATTTCCTGT	409
Qy	121	LeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValValLysLeuAlaThr	140
Db	410	CTGGCCCAACAGCCCGCTAGGAGTGTGGCCAGCCAGGTTGCTGTGCTCAAGCTTGCCACA	469
Qy	141	LeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThrAlaArg	160
Db	470	CTCGAAGACTTCTCTCTGCACCCCGAGTCCCAGATTGTGGAGGAGAACGGGACGACGCG	529
Qy	161	PheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAspGlnVal	180
Db	530	TTTGATGATGCCACACNAGGCGCTTCCAGCCCGCCATCATTTACTTGGGAANAGACACGGGTG	589
Qy	181	ThrValProGluGluProArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuAsp	200
Db	590	ACCGTGCCTGAGGAGTCCCGGCTCATCTCTTCCCAATGGCGTCTCCAGATCTTAGAT	649
Qy	201	ValGlnAspSerAspAlaGlySerTyrArgCysValAlaThrAsnSerAlaArgGlnArg	220
Db	650	GTCCAGGACAGTGATGACAGGCTCTTACCGCTGCGTGGCCACCAATTTCAGCCCGCCCAACGA	709

Qy	221	PheSerGlnGlnAlaSerLeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGly	240
Db	710	TTTACGACGAGGAGGCTCGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTACCAAGGGG	769
Qy	241	GlnAspValValIleValAlaAaProGluAsnThrThrValValSerGlyGlnAsnVal	260
Db	770	CAGATGTGTGTCATTGTGGCAGCCCAAGAGAACACCGGTAGTGTCTGGACAGAGTGTA	829
Qy	261	ValMetGluCysValAlaLeuSerAlaAspProThrProPheValSerTrpValArgGlnAsp	280
Db	830	GTGATGAGTGGTGGGCTCTCTGTGACCCCAACCCCTTTTGTGTCTGGTCCGACAGAT	889
Qy	281	GlyLysProLysThrAspValIleValLeuGlyArgThrAsnLeuLeuLeuAlaSer	300
Db	890	GGAAGGCTATCTCAGCGATGTCTGCTGAGCCCAACCCCTTTTGTGTCTGGTCCGACAGAT	949
Qy	301	AlaGlnProArgHisSerGlyValTyrValCysArgAlaAsnLeuProLeuThrArgAsp	320
Db	950	GCGCAGGCTCGGCACTCTGGAGTCTATGTCTGCCGAGCCCAACAGCCCGCCGACCGTGT	1009
Qy	321	PheAlaThrAlaAlaAlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaPro	340
Db	1010	TTGCCACTGCGGCTGTGAGCTCCGAGTGTCTGCTGCCCGCAGCCATCTCCGACGCGCCC	1069
Qy	341	GluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGlu	360
Db	1070	GAGCGGCTCTCGCGAGCGCGGCGCAGCACCGCGGCTCTGTTGTGCGCGGCTCCGCGGAG	1129
Qy	361	ProArgProAlaLeuHisTrpLeuHisAspGlyLeuProLeuArgProAsnGlyArgVal	380
Db	1130	CCACGGGCGCGCTGTCACCTGGCTGTCACGAGGATCCCGTTGCGACCAATGGGCGCGCTC	1189
Qy	381	LysValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGly	400
Db	1190	AGGTGTCAGGCGGTGGCGCAGGTTGTCTACTCATCTGATTCGGCTTCAGAGCGCTGCG	1249
Qy	401	TyrTyrGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaProLeuAla	420
Db	1250	TACTACAGTGTGTAGCAGAAAACAGCGCGGAACTGCTGTGTCGCTGCGCCCTGCGC	1309
Qy	421	ValValValArgGluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSer	440
Db	1310	GTAGTGTGTGCGCAGGGGCTGCCAGCGCCCGCTCGGGGTACAGCCACCGCTGAGC	1369
Qy	441	SerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGluGlnIleGly	460
Db	1370	AGCTCTCTGTGCTGTGGCTGGAGCGGCTGAGTTGTCACAGCAGCAATCATTTGGC	1429
Qy	461	PheSerLeuHisTyrGlnLysAlaArgValValAspAsnValGluTyrGlnPheAlaVal	480
Db	1430	TTCTCTCTTCACTACCAAAAGGCAAGGGAGTGGACAATGTGGAGTACCACTTTGCACTA	1489
Qy	481	AsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPhe	500
Db	1490	AACATGACACACAGAGCTGAGGTTCCGGACCTCGGAACCAACACACGATTTATGAGTTC	1549
Qy	501	TyrValValAlaTyrSerGlnLeuGlyValaSerArgThrSerSerProAlaLeuValHis	520
Db	1550	TACGTGTGGCTTACTCCAGCTGGGCGCAGCCGAACTCTCAGCCCGGCTGTGTGAT	1609
Qy	521	ThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSer	540
Db	1610	ACACTGGACGATGTTCACGCGCAGCCAGCTTACTTGTTCAGGCCCCCAACCCCTCG	1669
Qy	541	AspIleArgValAlaTrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLys	560
Db	1670	GACATCAGGTTGGATGCTGCTGCCCTCCAGCTGAGCAATGGACAGGTGCTGAG	1729
Qy	561	TyrLysIleGluTyrGlyLeuGlyLys---GluAspGlnValPheSerThrGluValPro	579
Db	1730	TACAAGATAGAGTACGGTTTGGGGAAGGAAGATCAGGTTTTCTCCACCGGAGTGCCT	1789

QY	580	GlyAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArg	599
DB	1790	GGAAATGAGACACAACTTACGTTAACTCACTTCAGCCAAACAAAGTGTACCGAGTCCGG	1849
QY	600	IleSerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThr	619
DB	1850	ATTTCAGCTGGCACTGGCGCTGGCTATGAGTGCCTTCTCAGTGGATGCAGACAGGACA	1909
QY	620	ProGlyValHisAsnGlnSerHisValProPheAlaProAlaGluLeuLysValArgAla	639
DB	1910	CTGGGTGTGCACAACAGAGCCATGTTCCCTTTGGCCCCCTGCAGAAATGGAAGGTGAGGGCA	1969
QY	640	LysMetGluSerLeuValValSerTrpGlnProProHisProThrGlnIleSerGly	659
DB	1970	AAGTGGAGTCCCTGGTGGTGTATGCGAGCCGCCCTCCACCCACCCAGATCTCTGGA	2029
QY	660	TyrLysLeuTyrTrpGlyGluValGlyThrGluGluAlaAspGlyAspArgProPro	679
DB	2030	TACAAACTCTACTGGAGAGAGTGGGAAACAGAGAGAGAGAGAGATGTTGACCGCCCCCA	2089
QY	680	GlyGlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLysLysValLys	699
DB	2090	GGGGTGTGTGAGATCAAGCTTGGGAGCTCGGGCCGCTGCGGTGGAAGAAAGTGAAG	2149
QY	700	GlnTyrGluLeuThrGlnLeuValProGlyValArgProTyrGluValLysLeuValAlaPhe	719
DB	2150	CAGTATGAATGACCCAGTTAGTCCCTGGGAGGCTGTACAGAGTGAAGCTGTGAGCTTTC	2209
QY	720	AsnLysHisGluAspGlyTyrAlaAlaValTrpLysGlyLysThrGluLysAlaProThr	739
DB	2210	AACAAACAGGAGAGCGCTAGCTGCTGTGTGGAAGGCAAGCGAGAGGCGCCACAG	2269
QY	740	ProAspLeuProIleGlnArgGlyProProLeuProProAlaHisValHisAlaGluSer	759
DB	2270	CCAGACCTGCTATCCAGAGGGGGCCACCGCTGCTCTGCCCATGTGTCCACGAGAGTCA	2329
QY	760	AsnSerSerThrSerIleTrpLeuArgTrpLysValProAspPheThrThrValLysIle	779
DB	2330	AACAGCTCCACTTCCATTGGCTTCGGTGGAGAGAGCCAGACTTTTACCAGTCAAGATT	2389
QY	780	ValAsnTyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTyr	799
DB	2390	GTCAACTACACTGTAGCTTTGGCCCTCGGGGCTCAGGAATGCTTCCCTGGTCACTAC	2449
QY	800	TyrThrSerSerGlyGluAspIleLeuIleGlyLeuLysProPheThrLysTyrGlu	819
DB	2450	TATACCAAGCTCTGAGAGAGACATTTCTATTGGGGCTGAAACCATTTTACCAAGTACGAG	2509
QY	820	PheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArg	839
DB	2510	TTTGGGTATACGTCACGAGTGGATATGATGGGGCTTTGGCTCCGTCGTAGAAAGCC	2569
QY	840	SerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThr	859
DB	2570	TCCACCTGTGCTGACCGGCTTCAACACCTCTTCTGACCTGCGCTGAGCCCCCTGACA	2629
QY	860	ProSerThrValArgLeuHisTrpCysProProThrGluProAsnGlyGluIleValGlu	879
DB	2630	CCATCCACCGTTGGTTACACTGGTGTCCCCCAGGAGCCCAATGGTGGATTTGTGGAG	2689
QY	880	TyrLeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThr	899
DB	2690	TATCTAATTTCTACAGCAACAACACACCGAGCCCCGAACACAGTGGACACTGCTCACC	2749
QY	900	ThrGluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhe	919
DB	2750	ACAGAGGGAAACATCTTCAGTGCAGAGGTCATGGCCCTAGAGGTGACACTCGGTATTTC	2809
QY	920	PheLysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspVal	939
DB	2810	TTCAAGATGGAGGCCGCACAGAGTGGGGCTGGGCCCTTTTCCCGCTTGCAGAGTGTG	2869
QY	940	IleThrLeuGlnGluThrPheSerAspSerLeuAspValHisAlaValThrGlyIleIle	959
DB	2870	ATTACTCTGCAAAAGACATTTCTCAGACTCTTGGATGTGCACCGCTCATCGGCAATCATC	2929
QY	960	ValGlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGln	979
DB	2930	GTGGGTGTCTGCTGGGCTTCTCTGCTCTCTGGCTGTGATGTGTGCTGCTGCTGCTGCTG	2989
QY	980	SerSerHisArgGluAlaLeuProGlyLeuSerSerSerGlyThrProGlyAsnProAla	999
DB	2990	AGCTCCACAGAGGAGCCCTGCCCGGATTTGCTCTCTCAGGCACCCAGGAACCCAGCG	3049
QY	1000	LeuTyrThrArgAlaArgLeuGlyProProSerValProAlaAlaHisGluLeuGluSer	1019
DB	3050	CTCTACACAGAGCTCGACTTGGGCCCCCAGTGTCCCTGCTGCCCATGAGTTGAGTCC	3109
QY	1020	LeuValHisProArgProGlnAspTrpSerProProSerAspValGluAspLysAla	1039
DB	3110	CTCGTGCACTCTCTGCTCCCGAGATTTGGTCCCCCACCCTCTCAGATTTGGAAGCAAGGCT	3169
QY	1040	GluValHisSerLeuMetGlyLysSerValSerAspCysArgGlyHisSerLysArgLys	1059
DB	3170	GAAGTACACAGCTTATGGTGGCAGTGTTCAGATTGCCGGGGCCACTCCAGAGAGAAAG	3229
QY	1060	IleSerTrpAlaGlnAlaGlyGlyProAsnTrpAlaGlySerTrpAlaGlyCysGluLeu	1079
DB	3230	ATCTCTGGGCTCAGGAGGGGACAAACTGGGCGAGGCTCTGGGCGAGGCTGTGAGCTG	3289
QY	1080	ProGlnGlySerGlyProArgProAlaLeuThrArgAlaLeuLeuProProAlaGlyThr	1099
DB	3290	CCCCAGGGTAGTGGTCCAGAGCGGCTCTGACCCCTGCTGCTCTGCTCTGCTCTGCTCTG	3349
QY	1100	GlyGlnThrLeuLeuLeuGlnAlaLeuValTyrAspGlyIleLysSerAsnGlyArgLys	1119
DB	3350	GGGCGAGACTCTCTCTGCTGACCCCTCTGATATATGACCCCTAAAGAGCAACCGGAGAAAG	3409
QY	1120	LysProSerProAlaCysArgAsnGlnValGluAlaGluValIleValHisSerAspPhe	1139
DB	3410	AAGCCGTCTCCAGCTGCGAGGAATCAGGTGGAAGCTGAGTCAATTGCTCCTCCTGCTCTG	3469
QY	1140	GlyAlaSerLysGlyCysProAspLeuHisLeuGlnAspLeuGluProGluProLeu	1159
DB	3470	GGTGCATCTCAAAAGGATGCTCTGACCTCCACCTCCCAAGACCTGGAGCCAGAGCAACTG	3529
QY	1160	ThrAlaGluThrLeuProSerThrSerGlyAlaValAspLeuSerGlnGlyAlaAspTrp	1179
DB	3530	ACTGCGAGACTCTGCTCTCCAGCTGTGAGCTGTGGATCTGTCTCTCAGAGGAGACTGG	3589
QY	1180	LeuGlyArgGluLeuGlyGlyCysGlnProThrThrSerGlyProGluArgLeuThrCys	1199
DB	3590	CTGGGCGAGGAGCTGGGAGGGTGCCCAACCAACAGTGGGCGAGAGGCTCACCTGC	3649
QY	1200	LeuProGluAlaAlaSerAlaSerCysSerCysSerAspLeuGlnProSerThrAlaIle	1219
DB	3650	TTGCCAAGAGAGCAGCAGTGCCTCTGCTCTGCTCAGACTCCAGCCAGCACTGCTATA	3709
QY	1220	GluGluAlaProGlyLysSerCysGlnProLysAlaLeuCysProLeuThrValSerPro	1239
DB	3710	GAGGAGGCCCTGGGAAAAGCTGCCAGCCCAAGCCCTGTGTCTCTTAACAGTCAGCCCA	3769
QY	1240	SerLeuProArgAlaProValSerSerAlaGlnValPro	1252
DB	3770	AGCCTTCCAGGGGCCCTGCTCTCTCTCTGCTCAGGTCCCC	3808
RESULT 6			
AKI22535			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
AKI22535			
Mus musculus mRNA for mKIAA1628 protein.			
AKI22535			
FLI_CDNA.			
Mus musculus (house mouse)			
Mus musculus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			

Db	1058	CGAGCCAAAGCCCGCAGCGGTGATTTCCGCCACTGCGGCTGCTGAGCTCCGAGTGCTT	1117
Qy	332	AlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAla	351
Db	1118	GCTGCCCGCAGCATCTCGCAGCGCCCGAGCGCTCTCGCGAGCGCGGCAGACCGCG	1177
Qy	352	ArgPheValCysArgAlaSerGlyGluProArgProAlaLeuHisTrpLeuHisAspGly	371
Db	1178	CGCTTCGTGTCCGGGGGTCCGGGAGCCACGCGCCCGCTGCACTGGCTGCACGCGGG	1237
Qy	372	IleProLeuArgProAsn-GlyArgValLysValGlnGlyGlyGlySerLeuValII	391
Db	1238	ATCCCGTTGGACCAATGGCGCGCTCAAGGTGCAGGGCGGTGGCGCAGCTTGGTCA	1297
Qy	391	eThrGlnIleGlyLeuGlnAspAlaGlyTyTrpGlnCysValAlaGluAsnSerAlaG1	411
Db	1298	CACCTCAGATGGCCCTTCAGGACGCTGGCTACTACAGTGGTACGAGAAACAGCGCGGG	1357
Qy	411	YThrAlaCysAlaAlaProLeuAlaValValArgGluGlyLeuProSerAlaPr	431
Db	1358	AACCTGCCTGTGCCGCTCGGCCCTCGCGGTAGTGGTGGCGAGGGGCTGCCAGCGCCC	1417
Qy	431	oThrArgValThrAlaThrProLeuSerSerSerSerValLeuValAlaTrpGluArgPr	451
Db	1418	GACTCGGTGCACGCCACGCCGCTGAGCAGCTCCTCTGTGTGTGGCTGGAGCGGCC	1477
Qy	451	oGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTyGlnLysAlaArgGlyVa	471
Db	1478	TGAGTTGCACAGCAGCAATCATTTGGCTTCTCTCTACCTACCAAAAGGCAAGGGAGT	1537
Qy	471	LaAspAsnValGluTyTrpGlnPheAlaValAsnAsnAspThrThrGluLeuGlnValArgAs	491
Db	1538	GGACAAATGTGAGTACCAATTTGCAATTAACATATGACACACAGAGCTGCAGGTTCCGGA	1597
Qy	491	pLeuGluProAsnThrAspTyTrpGluPheTyValValAlaTySerGlnLeuGlyAlaSe	511
Db	1598	CCTGGAACCCCAACAGCATATGAGTTCTACGTGTGGCTTCTCCAGCTGGGGCCAG	1657
Qy	511	rArgThrSerSerProAlaLeuValHisThrLeuAspAspValProSerAlaAlaProG1	531
Db	1658	CCGAACCTCCAGCCCGCCCTGGTGATACACTGGAGATGTCCTCCAGCGAGACCCCA	1717
Qy	531	nLeuThrLeuSerSerProAsnProSerAspIleArgValAlaTrpLeuProLeuProSe	551
Db	1718	GCTTACCTTGTCCAGCCCAACCCCTCGGACATCAGGGTGGCATGGCTGCCCTGCCCTC	1777
Qy	551	rSerLeuSerAsnGlyGlnValLeuLysTyLysIleGluTyGlyLeuGlyLys---	570
Db	1778	CAGCTGAGCAATGGACAGAGTGTCTGAAGTACAAGATAGAGTACCGTTTGGGGAAGGA	1837
Qy	570	uAspGlnValPheSerThrGluValProGlyAsnGluThrGlnLeuThrLeuAsnSerLe	590
Db	1838	AGATCAGGTTTCTCCACCGAGGTGCCTGGAAATGAGACACAACTTACGTTAAACTCACT	1897
Qy	590	uGlnProAsnLysValTyArgValArgIleSerAlaGlyThrGlyAlaGlyTyGlyVa	610
Db	1898	TCAGCCAAACAAAGTGTACCGAGTCCGGAATTCAGCTGGCACTGGCGCTGGCATGGAT	1957
Qy	610	lProSerGlnTrpMetGlnHisArgThrProGlyValHisAsnGlnSerHisValProPh	630
Db	1958	CCCTTCTCAGTGGATGCAGCAGACGACACCTGGTGTGCACAAACAGAGCCATGTTCCTT	2017
Qy	630	eAlaProAlaGluLeuLysValArgAlaLysMetGluSerLeuValValSerTrpGlnPr	650
Db	2018	TGCCCTCTCAGAAATGAAGGTGAGGGCAAGATGGAGTCCCTGGTGGTGTCAATGGCAGCC	2077
Qy	650	oProProHisProThrGlnIleSerGlyTyTrpLysLeuTyTrpGlyGluValGlyThrG1	670
Db	2078	GGCCCTCACCACCCAGATCTCTGGATACAAACTCTACTTGGAGAGGTTGGGAAACAGA	2137
Qy	670	uGluAlaAlaAspGlyAspArgProProGlyGlyArgGlyAspGlnAlaTrpAspValG1	690
Db	2138	GGAGGAGCAGATGGTGTAGCCGCCCGCCAGGGGGTGTGGAGATCAAGCTTGGGACGTCCG	2197
Qy	690	yProValArgLeuLysLysValLysGlnTyTrpGluLeuThrGlnLeuValProGlyAr	710
Db	2198	GCCCGTGGCGCTGAAGAGAAAGTGAAGCAGATATGAAGTATGAAGTATGATCCCTGGCAG	2257
Qy	710	gProTyGluValLysLeuValAlaPheAsnLysHisGluAspGlyTyAlaAlaValTr	730
Db	2258	GCTGTACGAGGTGAAGCTCGTAGCTTCAACAACACAGAGGACGGCTACGCTGCTGTGTG	2317
Qy	730	plysGlyLysThrGluLysAlaProThrProAspLeuProIleGlnArgGlyProProLe	750
Db	2318	GAAGGCAAGACGGAGAGGCGCCACGACAGCTGCTATCCAGAGGGGCGCCACCGCT	2377
Qy	750	uProProAlaHisValHisAlaGluSerAsnSerSerThrSerIleTrpLeuArgTrpLy	770
Db	2378	GCCTCTGCCCATGTCCACGACAGGTCAAAACAGCTCCACTTCCATTTGGCTTGGTGGAA	2437
Qy	770	sLysProAspPheThrThrValLysIleValAsnTyThrValArgPheGlyProTrpG1	790
Db	2438	GAAGCCAGACTTTTACCACTGTCAAGATTGTCAACTACACTGTATCGCTTCGGGCCCTGGGG	2497
Qy	790	yLeuArgAsnAlaSerLeuValTyTrpTyThrSerSerGlyGluAspIleLeuIleG1	810
Db	2498	GCTCAGGAATGCTTCCCTGGTCACTTATACCAAGCTCTGGAGAGACATTTCTCATTTGG	2557
Qy	810	yGlyLeuLysProPheThrLysTyxGluPheAlaValGlnSerHisGlyValAspMetAs	830
Db	2558	CGGCTCGAAACCATTTACCAAGTACGAGTTTGGGTGACAGTCCCAACGAGTGGATATGA	2617
Qy	830	pGlyProPheGlySerValValGluArgSerThrLeuProAspArgProSerThrProPr	850
Db	2618	TGGGCCCTTGGCTCGCTCGTAGAACCGCTCCACCTGCTGACCGGCCCTTCAACACCTCC	2677
Qy	850	oSerAspLeuArgLeuSerProLeuThrProSerThrValArgLeuHisTrpCysProPr	870
Db	2678	TTCTGACCTGCGCTGAGCCCTGACACCATCCACCGTTCGGTTACACTGGTGTCCCC	2737
Qy	870	oThrGluProAsnGlyGluIleValGluTyTrpLeuLysLeuTySerAsnAsnHisThrG1	890
Db	2738	CACGAGCCCAATGGTGAGATTGTGGAGTATCTATTTCTTACAGCAACACACACACCCA	2797
Qy	890	nProGluHisGlnTrpThrLeuLeuThrThrGluGlyAsnIlePheSerAlaGluValHi	910
Db	2798	GCCCGAACACCACTGAGTGACACTCTACCAACAGAGGGGAAACATCTTCACTGACAGGTCCA	2857
Qy	910	sGlyLeuGluSerAspThrArgTyxPhePheLysMetGlyAlaArgThrGluValGlyPr	930
Db	2858	TGGCCTTAGAGAGTGAACACTCGGTATTCTTCAAGATGGGAGCCCGCACAGAGTGGGGCC	2917
Qy	930	oGlyProPheSerArgLeuGlnAspValIleThrLeuGlnGluThrPheSerAspSerLe	950
Db	2918	TGGGCCCTTTTCCCGCTTCAGGATGTGATTACTCTGCAAAAGACATTTCTCAGACTCCTT	2977
Qy	950	uAspValHisAlaValThrGlyIleIleValGlyValCysLeuGlyLeuLeuCysLeuLe	970
Db	2978	GGATGTGCACCGCTCACGGGCATCATCGTGGGTGTCTGCTGGGCGCTTCTCTGCTCCT	3037
Qy	970	uAlaCysMetCysAlaGlyLeuArgGlnSerSerHisArgGluAlaLeuProGlyLeuSe	990
Db	3038	GGCTTCGATGTGTGCTGGCCCTACGACGAAGCTCCCAAGAGAGCCCTGCCCCGATGTGC	3097
Qy	990	rSerSerGlyThrProGlyAsnProAlaLeuTyThrArgAlaArgLeuGlyProProSe	1010
Db	3098	CTCTCAGGCACCCAGGAAACCCAGCGCTCTACACAGAGCTCGACTTGGGCCCCCAG	3157
Qy	1010	rValProAlaAlaHisGluLeuGluSerLeuValHisProArgProGlnAspTrpSerPr	1030
Db	3158	TGTCCCTGCTGCCCATAGTTGGAGTCCCTCGTGCACTCTCGTCCCGCAGGATTTGGTCCC	3217
Qy	1030	oProProSerAspValGluAspLysAlaGluValHisSerLeuMetGlyGlySerValSe	1050
Db	3218	ACCACCTCAGATGTGGAGACAAAGCTGAAGTATGAGCTTATGGTGGCAGTGTTC	3277

QY 1050 rAspCysArgGlyHisSerLysArgLysIleSerTrpAlaGlnAlaGlyGlyProAsnTr 1070
 DB 3278 AGATTGCCGGGCCCACTCAAGAGAAAGATCTCTCGGCTCAGGCGGGGACCAAACTG 3337
 QY 1070 pAlaGlySerTrpAlaGlyCysGluLeuProGlnGlySerGlyProArgProAlaLeuTh 1090
 DB 3338 GCGAGCTCTCGGGCAGGCTGTGAGCTGCCAGGGTAGTGTCACAGCGCGCTCTGAC 3397
 QY 1090 rArgAlaLeuLeuProProAlaGlyThrGlyGlnThrLeuLeuLeuGlnAlaLeuValTy 1110
 DB 3398 CGGTGCTCTGCTGCTCTCAGCGGGAACCGGGCAGACACTGCTGCTGCAAGCCCTGGTATA 3457
 QY 1110 rAspGlyIleLysSerAsnGlyArgLysIleProSerProAlaCysArgAsnGlnValGl 1130
 DB 3458 TGACGCCATAAGACCAACGGGAGAAAGACCGTCCCGAGCTGCGAGGAATCAGGTGGA 3517
 QY 1130 uAlaGluValIleValHisSerAspPheGlyAlaSerLysGlyCysProAspLeuHisLe 1150
 DB 3518 AGCTGAGGTCAATGTGCCACTCGACTTCGGTGTCATCCAAAGGATGTCTGACCTCCACT 3577
 QY 1150 uGlnAspLeuGluProGluProLeuThrAlaGluThrLeuProSerThrSerGlyAl 1170
 DB 3578 CCAAGACTCGAGCCAGGAGCACTGACTGCGAGAGACTCTGCCCTTCACGCTCGGAGC 3637
 QY 1170 aValAspLeuSerGlnGlyAlaAspTrpLeuGlyArgGluLeuGlyCysGlnProTh 1190
 DB 3638 TGTGATCTGCTCAAGAGCAGACAGCTGCTGGCAGGAGCTGGGAGGCTGCCAACAC 3697
 QY 1190 rThrSerGlyProGluArgLeuThrCysLeuProGluAlaAlaSerAlaSerCysSerCy 1210
 DB 3698 AACCAAGTGGCCAGAGAGGCTCACCTGCTGTCAGAGAGCAGTGCCTCTGCTGCTG 3757
 QY 1210 eSerAspLeuGlnProSerThrAlaIleGluAlaProGlyLysSerCysGlnProLy 1230
 DB 3758 CTCAGACTCCAGCCAGCAGCTGCTATAGAGAGGCCCTCGGAAAGCTCCAGGCCCA 3817
 QY 1230 sAlaLeuCysProLeuThrValSerProSerLeuProArgAlaProValSerSerAlaGl 1250
 DB 3818 AGCCCTGTGCTCTTAAGCAGTCAGCCAGCCAGCTCCAGGAGGCCCTGCTCTCTGCTCA 3877
 QY 1250 nValPro 1252
 DB 3878 GGTCCCC 3884

RESULT 7
 AB052622
 LOCUS Homo sapiens hDDM36 mRNA, complete cds.
 DEFINITION
 ACCESSION AB052622
 VERSION AB052622.1 GI:19570397
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 AUTHORS Murakami, H., Nakamura, T., Nakayama, T., Yamamoto, H., Hosaka, T.,
 Aoyama, T., Nagayama, S., Oka, M., Kiyono, T., Sasaki, M.S., Nakamura, T.
 and Toguchida, J.
 TITLE Up-regulation of a ras effector and down-regulation of a cell
 adhesion molecule are associated with transformation of osteoblasts
 unpublished
 JOURNAL
 REFERENCE
 AUTHORS Toguchida, J., Nakamura, T., Murakami, H., Nakayama, T. and Nakamura, T.
 TITLE Direct Submission
 JOURNAL
 Submitted (11-DEC-2000) Junya Toguchida, Kyoto University,
 Institute for Frontier Medical Sciences; 53 Kawahara-cho, Shogoin,
 Sakyo-ku, Kyoto, Kyoto 606-8507, Japan
 (E-mail: togjun@frontier.kyoto-u.ac.jp, Tel: 81-75-751-4134,
 Fax: 81-75-751-4144)
 FEATURES
 Location/Qualifiers
 1..6485
 /organism="Homo sapiens"

/mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="15"
 /clone="hDDM36"
 /cell_line="NH08t"
 /cell_type="osteoblast"
 /notes="similar to mouse DDM36 and Nope
 (NM_020043)-vector:phDDM36"
 1..6485
 /gene="hDDM36"
 211..3963
 /gene="hDDM36"
 /notes="similar to mouse DDM36 and Nope (NP_064427)"
 /codon_start=1
 /evidence=experimental
 /product="hDDM36"
 /protein_id="BA886306.1"
 /db_xref="GI:19570398"
 /translation="MARGDAGRGRGLLALFCLLAARGELLFQETTVELSCGVGPIQ
 VIILGPEQAAVLNCSLGAAGAAAPPTVTWSKQDITLLEHDLHLLPNSLSLWOPFAPN
 GDSVPEAVGVIRGNYSCLAHGPLVGLAVKLAATLADLSLHPESOTVSENGTA
 RFCHLEGLPAPIITWEKQVTLPEEPRLJLVNGLQILDLQVSDAGPFRVATNSA
 RQHSQALLSVAHRSGLASTRGDQVIVAAPENTVVGQSVVMECVASADPTFPVS
 WVRQDKPISTDIVILGRITNLLIANAQFVHSVYVCRANKPRTDFATNAALFVLA
 PAITQPEALSRSTRASTARFVCRASGEPRALRWLHNGAPLRNGRVKVGQGGSLVI
 TOIGLDAGTYQCVAENAGMACAAASLAVVREGLPSA PRTVATPLSSASVLVAME
 RPEHSSQIIGFSLHYOKARGMONVEYQPAVNDTTELOVRDLEPNDTYEFYVAYSQ
 LGSRTSTPALVHTLDDVPSAQLSLSPNSDIRVAMLPPLPSLNGQVVKYKIEY
 GLGKEDIPTSEVRNETQMLNSLQPNKYRVRISAGTAPGAFGAPSGMHHRTFPMH
 RQSHVPPAPAEKVQAKMSLVSWQPPHPTQISGKYWRVEYGAEEANGDLPGG
 RQDAQMDVGVRLKKYKQYELTQVLPRLYEVKLVAFNKHEDDGYAAVMKGTAKPA
 PMP10RGPPPLPAHVHAESENSTSIWLRWKKPDPPTVKIIVNVTSPWGLRNASLV
 TYTSSGDIIGLGLKPTKYBPAVOSHGVMDPGVSVVERSTLPDRBPTPPSDLR
 SPLTPSVRLHWCPTPEPNGEIVEYLLIYSNHTQPEHQTLLTQGNISAEVHGLE
 SDTRYFPMGARTGVPPFSLQDVITLQKLSDSLDMHVSVTGIIIVGVCLLCLLA
 CMAGLRSRPHRESLPGLSSTATFGNPAVYSLRGLPSPPPAAHELSLHVPQDMS
 PPSVEDRPAEVLHSLGGVSEGRSHSKRIKISWAPSGLSWAGSACCTGAGPQAPRA
 LTRALLPAGTGTLLIQLAVYDAIKNGRKKKPPACRNOVEARVIVHSDFASNGNP
 DLHLQDLEPDLPPRAPDLISGVDPGQCAWLDRELGCCLAAAPGPDRLTCLPEAA
 SASCSYDLPDQGVLESTPQDSQLKSPCLGASPLPRFSPVSSA"

BASE COUNT 1293 a 1926 c 1772 g 1494 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2,87e-225 Length: 6485
 Score: 5746.00 Matches: 1098
 Percent Similarity: 92.09% Conservative: 54
 Best Local Similarity: 87.77% Mismatches: 96
 Query Match: 87.01% Indels: 5
 DB: 9 Gaps: 2

US-09-754-997A-2 (1-1252) x AB052622 (1-6485)

QY 1 MetAlaArgAlaAspThr-----GlyArgGlyLeuLeuValLeuThrPheCysLeuLeu 18
 DB 211 ATGCGCGGGGGAGCGCGCGCGCGCGCGCTCTCGCGTTGACCTCTCTGCTGTTG 270
 QY 19 SerAlaArgGlyGluLeuLeuProLeuProGlnGluThrThrValLysLeuSerCysAspGlu 38
 DB 271 GCCGCGCGGGGAGCTGCTGTTGCCCGCAGAGACGACTGTGGAGCTGAGCTGTGGAGTG 330
 QY 39 GlyProLeuGlnValIleLeuGlyProGluGlnAlaValValLeuAAspCysThrLeuGly 58
 DB 331 GGGCCACTGCAAGATGATCTCGGGCCAGAGAGGCTGACGTCTAACTGTAGCTGTGGG 390
 QY 59 AlaThrAlaAlaGlyProProThrArgValThrTrpSerLysAspGlyAspThrValLeu 78
 DB 391 GCTGCTGCCCTGGACCCCCCACCAGCGGTGACCTGGAGACAGATGGGACACCTGCTG 450
 QY 79 GluHisGluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGlu 98
 DB 451 GAGCAGCACCACTTACACCTGCTGCCCAATGGTTCCCTGTGGCTGTGCCAGCCTAGCA 510

Db 2668 GAGTTTGCAGTCAGTCTCAAGCGCTGGACATGATGGCCCTTTCCGGCTCTGTGGTGAG 2727
 Qy 839 ArgSerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeu 858
 Db 2728 CGCTCCACCTCGCTGACCGGCGCTCCACACCCCATCCGACCTCGAGCTGAGCCCTCG 2787
 Qy 859 ThrProSerThrValArgLeuHisThrProProProProProProProProProProPro 878
 Db 2788 ACACCGTCCAGCGTTCCGCTGACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2847
 Qy 879 GluThrLeuLeuLeuSerAsnAsnHisThrGlnProGluHisGlnThrThrLeuLeu 898
 Db 2848 GAGTATCTGATCTGTACAGCAGCAACCAACAGCAGCTGAGCAGCAGTGGACCTTGCTC 2907
 Qy 899 ThrThrGluGluValPheSerAlaGluValHisGlyLeuGluSerAspThrArgTyr 918
 Db 2908 ACCACGACGAGGAACATCTTCAGTCTGAGGTCCATGCTGGAGCGACACTCGGTAC 2967
 Qy 919 PhePheLeuMetGlyValArgThrGluValGlyProGlyProPheSerArgLeuGlnAsp 938
 Db 2968 TTCTTCAAGATGGGGCGCGCACAGAGGTGGGACCTGGGCGCTTTCTCCGCGCTGCAGGAT 3027
 Qy 939 ValIleThrLeuGlnGluThrPheSerAspSerLeuAspValHisAlaValThrGlyIle 958
 Db 3028 GTGATCAGCTCCAGAGAGAGCTGTACACTCGCTGGACATGCACTCAGTCACGGGCATC 3087
 Qy 959 IleValGlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArg 978
 Db 3088 ATCGTGGGTGTCTGCTGGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3147
 Qy 979 GlnSerSerHisArgGluAlaLeuProGlyLeuSerSerSerGlyThrProGlyAsnPro 998
 Db 3148 CGCAGCCCGCACAGGGATCCCTCCAGCGCTGTCTCCAGCGCTGTCTCCAGCGCTGTCTCCAG 3207
 Qy 999 AlaLeuThrThrArgAlaArgLeuGlyProProSerValProAlaAlaHisGluLeuGlu 1018
 Db 3208 GCGCTGTACTCAGAGCTGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3267
 Qy 1019 SerLeuValHisProArgProGlnAspTrpSerProProProSerAspValGluAspLys 1038
 Db 3268 TCCCTTGTGACCCCGACCCAGGACTGTGTCCCGCGCACCTCAGACGTGGAGGACAGG 3327
 Qy 1039 AlaGluValHisSerLeuMetGlyGlySerValSerAspCysArgGlyHisSerLysArg 1058
 Db 3328 GCTGAAGTGACAGCGCTTATGGGTGGCGGTGTTCTGAAGGCGCGAGTCACTCCAAAGA 3387
 Qy 1059 LysIleSerTrpAlaGlnAlaGlyCysProAsnTrpAlaGlySerTrpAlaGlyCysGlu 1078
 Db 3388 AAGATCTCTGGGCTCAACCAAGCGGCTGAGCTGGGCTGTCTTCCGGCGAGGCTGTGAG 3447
 Qy 1079 LeuProGlnGlySerGlyProArgProAlaLeuThrArgAlaLeuLeuProProAlaGly 1098
 Db 3448 CTGCCCCAG---GCAAGGCGCGCGCGCGCTCTGACCCCGCGCGCTGCTGCTGCTGCTG 3504
 Qy 1099 ThrGlyGlnThrLeuLeuGlnAlaLeuValTyrAspGlyIleLysSerAsnGlyArg 1118
 Db 3504 ACTGGGCGACGCTGTCTGAGGCTCTGGGTGTACGACGCCATAAAGGCGCAATGGGAGG 3564
 Qy 1119 LysLysProSerProAlaCysArgAsnGlnValGluAlaGluValIleValHisSerAsp 1138
 Db 3565 AAGAAAGTCAACCCCGAGCTGACGAGAACGAGGTGGAGGCTGAAGTCAATGTCCACTCTGAC 3624
 Qy 1139 PheGlyAlaSerLysGlyCysProAspLeuHisLeuGlnAspLeuGluProGluGluPro 1158
 Db 3625 TTTAGTGCATCTAACCGGAACCTGACCTCCATCTCCAGACCTGGAGCCTGGAGACCC 3684
 Qy 1159 LeuThrAlaGluThrLeuProSerThrSerGlyAlaValAspLeuSerGlnGlyAlaAsp 1178
 Db 3685 CTGCTCCAGAGCTCTGATCTCATCTCGGGTGTGGGGATCCAGGCGAGGGGCGACCC 3744
 Qy 1179 TrpLeuGlyArgGluLeuGlyCysGlnProThrThrSerGlyProGluArgLeuThr 1198
 Db 3745 TGGCTGGACAGGAGTTGGAGGGGTGTGAGCTGGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 3804

Qy 1199 CysLeuProGluAlaAlaSerAlaSerCysSerCysSerAspLeuGlnProSerThrAla 1218
 Db 3805 TGCTTGCAGAGGCGAGCAGTGTCT 3864
 Qy 1219 IleGluGluAlaProGlyLysSerCysGlnProLysAlaLeuCysProLeuThrValSer 1238
 Db 3865 CTAGAGAGACCCCTGGAGATAGCTGCCAGTCAAAATCCCTCTCTCTCTCTCTCTCTCTCTCT 3924
 Qy 1239 ProSerLeuProArgAlaProValSerSerAla 1249
 Db 3925 CCAGGCTGCCAGATCCCGGTCT 3957
 RESULT 8
 AX418406 3741 bp DNA linear PAT 18-JUN-2002
 LOCUS AX418406
 DEFINITION Sequence 1 from Patent WO0206329.
 ACCESSION AX418406
 VERSION AX418406.1 GI:21523336
 KEYWORDS
 SOURCE unidentified
 ORGANISM unidentified
 1 unclassified.
 REFERENCE
 AUTHORS Rastelli, L., Shinkets, R.A., Zerhusen, B., Malyanar, U.M. and Padigaru, M.
 TITLE Human polynucleotides and polypeptides encoded thereby
 JOURNAL Patent: WO 0206329-A 1 24-JAN-2002;
 Curagen Corporation (US)
 FEATURES
 source
 1. 3741
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 /note="NOVX nucleic acid with homology to NOPE-like protein"
 BASE COUNT 708 a 1209 c 1150 g 673 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 5,38e-220 Length: 3741
 Score: 5610.00 Matches: 1086
 Percent Similarity: 91.07% Conservative: 56
 Best Local Similarity: 86.60% Mismatches: 101
 Query Match: 84.95% Indels: 13
 DB: 6 Gaps: 8
 US-09-754-997A-2 (1-1252) x AX418406 (1-3741):
 Qy 1 MetAlaArgAlaAspThr-----GlyArgGlyLeuLeuValLeuThrPheCysLeuLeu 18
 Db 1 ATGGCGGGGGGAGCG 60
 Qy 19 SerAlaArgGlyGluLeuProLeuProGlnGluThrThrValLysLeuSerCysAspGlu 38
 Db 61 GCCCGCGCGGGAGCTGTGTTCGCCAGAGACGACTGTGGAGCTGAGCTGTGGAGTG 120
 Qy 39 GlyProLeuGlnValIleLeuGlyProGluGlnAlaValValLeuLeuAspCysThrLeuGly 58
 Db 121 GGGCCACTGCAAGTGATCTCTGGGCGCGAGCGAGCTGAGTGTAACTGTAGCTGTGGG 180
 Qy 59 AlaThrAlaAlaGlyProProThrArgValThrTrpSerLysAspGlyAspThrValLeu 78
 Db 181 GCTGCTGCGCTGGACCCCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
 Qy 79 GluHisGluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGlu 98
 Db 241 GAGCAGCAGCACTTACACTGTCTGGCCAAATGTTCCCTCTGTGGCTGTCCAGCCACTAGCA 300
 Qy 99 GlnGluAspSerAspAspGluGluAlaLeuArgIleTrpLysValThrGluGlySerTyr 118
 Db 301 CCCAATGGCAGTGCAGAGT--CAGTCCCTGAGGCTGTGGGG-GTCAATTGAAGGCACTAT 357

QY	62	AlaGlyProProThrArgValThrTrpSerLysAspGlyAspThrValLeuGluHisGlu	81	DB	1201	GTGGTGGCGAGGGGCTGCCAGCGCCCGACTCGGGTCAACAGCCAGCGCGCTGAGCAGC	1260
DB	121	GCTGGGCGCTCCGACAGGCTGACATGGAGCAAGATGGAGACACTGTACTAGACATGAG	180	QY	442	SerSerValLeuValAlaTrpGluArgProGluLeuHisSerGluGlnIleLeuGlyPhe	461
QY	82	AsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGluGlnGluAsp	101	DB	1261	TCCTCTGTGCTGGTGGGCGGCTGGAGCGGCTGAGTTGCACAGCGAGCAATCATTTGGCTTC	1320
DB	181	AACCTGCACCTGTACCAATGGCTCCCTGTGGCTGTCTCTCAACCCCTAGAGCAAGAAGAC	240	QY	462	SerLeuHisTrpGlnLysAlaArgGlyValAspAsnValGluTrpGlnPheAlaValAsn	481
QY	102	SerAspAspGluGluAlaLeuArgIleTrpLysValThrGluGlySerTrpSerCysLeu	121	DB	1321	TCCTCTCACTACCAAAAGGCAAGGAGTGGACATGTGGAGTACCAAGTTTGCAGTAAAC	1380
DB	241	AGCGATGATGAGGAAGCTCTTAGGATCTGGAAGTCACTGAGGCGAGCTATTCTCTGTCG	300	QY	482	AsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTrpGluPheTrp	501
QY	122	AlaHisSerProLeuGlyValValAlaSerGlnValAlaValValLysLeuAlaThrLeu	141	DB	1381	AATGACACACAGAGCTCGAGTTTCGGACCTCGGAACCCCAACACGATTTATGAGTTCTAC	1440
DB	301	GCCCCACAGCCCGTAGGAGTGGTGGCCAGCCAGGTTGCTGTGGTCAAGCTTGCACACTC	360	QY	502	ValValAlaTrpSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThr	521
QY	142	GluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThrAlaArgPhe	161	DB	1441	GTGGTGGGCTACTCCAGCTGGGGCCAGCGCAACTCCAGCCAGCCCTGGTGATACA	1500
DB	361	GAAGACTTCTCTGACCCCGAGTCCAGATTTGTGAGAGAGAAACGGGACAGCACGCTTT	420	QY	522	LeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAsp	541
QY	162	GluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAspGlnValThr	181	DB	1501	CTGGAGCATGTCCCGAGCGCAGCACCCAGCTTACTTGTCCAGCCCAACCCCTCGGAC	1560
DB	421	GAATGCCACACCAAGGGCTTCAGGCCCCCATCATTTCTGGGAAGGACCAAGGTGACC	480	QY	542	IleArgValAlaTrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLysTrp	561
QY	182	ValProGluGluProArgLeuLeuThrLeuProLysTrpLeuLeuGlnIleLeuAspVal	201	DB	1561	ATCAGGGTGGCATGCTGCCCTGCCCTCCAGCTGAGCAATGGACAGGTGCTGAAGTAC	1620
DB	481	GTGCTGAGGAGGAGCCCGGCTCATCTCTCCCAAGTGGCTCTCCAGATCTAGATGTC	540	QY	562	LysIleGluTrpGlyLeuGlyLysGluAspGlnValPheSerThrGluValProGlyAsn	581
QY	202	GlnAspSerAspAlaGlySerTrpArgCysValAlaThrAsnSerAlaArgGlnArgPhe	221	DB	1621	AAGATAGAGTACGGTTTGGGGAAGGAAGATCAGGTTTCTCCAGCCAGGTGCTGGAAT	1680
DB	541	CAGGACAGTATGATCAGGCTCTACCGCTGGTGGCCACCATAATTCAGCCCGCCAAAGATT	600	QY	582	GluThrGlnLeuLeuLeuAsnSerLeuGlnProAsnLysValTrpArgValArgIleSer	601
QY	222	SerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGln	241	DB	1681	GAGACACAACTTACGTTAACTCATTTCAGCCAAACAAGTGTACCGAGTCCGAGTTTCA	1740
DB	601	AGCCAGGAGGCGCTCGCTACTGTGGCCCTCAGAGGCTCTTTGAGGCTACCAAGGGGCG	660	QY	602	AlaGlyThrGlyValaGlyTrpGlyValProSerGlnTrpMetGlnHisArgThrProGly	621
QY	242	AspValValIleValAlaProGluAsnThrThrValValSerGlyGlnAsnValVal	261	DB	1741	GCTGGCACTGGCGCTGGCTATGAGTCCCTTCTCAGTGGATGCAGCACAGGACACCTGGT	1800
DB	661	GATGTGTCTATTGGCAGCCCGAGAGAACACCGTAGTGTCTGACAGCAAGATGTAGTG	720	QY	622	ValHisAsnGlnSerHisValProPheAlaProAlaGluLeuLysValArgAlaLysMet	641
QY	262	MetGluCysValAlaSerAlaAspProThrProPheValSerTrpValArgGlnAspGly	281	DB	1801	GTGCACAACAGAGCATGTTCCTTGTCCCTTCAGCAATTTGAAGTGTAGGGCAAGATG	1860
DB	721	ATGGAGTGGCTGGCTCTGCTGACCCACCACCCCTTTGTCTCTGGGTCCGACAGGATGA	780	QY	642	GluSerLeuValValSerTrpGlnProProHisProThrGlnIleSerGlyTrpLys	661
QY	282	LysProIleSerThrAspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSerAla	301	DB	1861	GAGTCCCTGGTGTGTCATGGCAGCCGCCCTCACCCACCCAGATCTCTGATACAA	1920
DB	781	AAGCTATCTCCACGATGTATCTGTGGGCGGACCAATCTACTCATCGCAGCGCG	840	QY	662	LeuTrpTrpGlyGluValGlyThrGluGluGluAlaAspGlyAspArgProGlyGly	681
QY	302	GlnProArgHisSerGlyValTrpValCysArgAlaAsnLysProLeuThrArgAspPhe	321	DB	1921	CTCTACTGGGAGAGTGGGAACAGAGGAGGAGGAGATGGTGACCCGCCAGGGGGT	1980
DB	841	CAGCCTCGGCATCTGGAGTCTATGTCTGCGAGGCCAACAGCCCTCTACCGGTGACTTC	900	QY	682	ArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLysLysLysValLysGlnTrp	701
QY	322	AlaThrAlaAlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGlu	341	DB	1981	CGTGGAGATCAAGCTTGGACGTCGGGCGCGTGGCGCTGAGAGAAAGTGAAGCAGTAT	2040
DB	901	GCCACTGGCGCTGCTGAGCTCCGAGTCTGCTGTGCTGCCACGCCATCTCGCAGGACCCGAG	960	QY	702	GluLeuThrGlnLeuValProGlyArgProTrpGluValLysLeuValAlaPheAsnLys	721
QY	342	AlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluPro	361	DB	2041	GAACTGACCACTGTAGTCCCTGGCAGCGCTGACAGGTGAAGCTCTGATCTTCAACAA	2100
DB	961	GCGCTCTCGCGGACCGCGGCGCAGCACCGCGCTTCTGTGTGCGGCGCTCCGCGGAGCCA	1020	QY	722	HisGluAspGlyTrpAlaAlaValTrpLysGlyLysThrGluLysAlaProThrProAsp	741
QY	362	ArgProAlaLeuHisTrpLeuHisAspGlyLysProLeuArgProAsnGlyArgValLys	381	DB	2101	CACGAGGACGGTACGCTGTGTGGAAAGGGGAGAGCGGAGGAGGCGGCCCGCCAGAC	2160
DB	1021	CGGCCCGCGCTGACTGCTGCGACGAGCGGATCCGTTGCGACCCCAATGGGCGCGTCAAG	1080	QY	742	LeuProIleGlnArgGlyProProLeuProProAlaHisValHisAlaGluSerAsnSer	761
QY	382	ValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTrp	401	DB	2161	CTGCCCTATCCAGAGGGGCGCACCGCTGCTCTCTGCCCATGTCCACCGAGGTCAACAGC	2220
DB	1081	GTGACGGGCGGTGGCGGACGCTGGTCACTCATCAGATCGGCTGCGAGCGCTGGCTAC	1140	QY	762	SerThrSerIleTrpLeuArgTrpLysProAspPheThrThrValLysIleValAsn	781
QY	402	TyrGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaProLeuAlaVal	421	DB	2221	TCCACTTCCATTTGGCTTCGGTGGAGAACCCAGACTTTTACCCTGCTCAAGATTGTCAAC	2280
DB	1141	TACCAGTGGCTAGCAGAAAACAGCGCGGAACCTGCCGTGCCGCTCGCGCGGTA	1200	QY	782	TyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTrpTrpThr	801
QY	422	ValValArgGluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSer	441				

Db 2281 TACACTGTACGCTTCGGCCCTCGGGGCTCAGGAATGCTTCCTCGTCACTACTATACC 2340
 QY 802 SerSerGlyuAspIleuLeuGlyGlyLeuLysProPheThrLysThrGluPheAla 821
 Db 2341 AGCTCTGGAGAAGACATCTCATTTGGCGCCCTGAAACCATTTTACCAAGTACGAGTTTGG 2400
 QY 822 ValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSerThr 841
 Db 2401 GTACAGTCCCAACGAGTGGATATGGATGGCCCTTTGGCTCGCTCGTAGAAGCGCTCCACC 2460
 QY 842 LeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrProSer 861
 Db 2461 CTGCTGACCGCCCTTCAACACTCTCTTGACCTGCGCCTGAGCCCTGACACCATCC 2520
 QY 862 ThrValArgLeuHisThrCysProProThrGluProAsnGlyGluIleValGluThrLeu 881
 Db 2521 ACCGTTGCGTTACACTGTTGTCCTCCCGCCCGAGGCCAATGGTACAGTTGTGAGATCTA 2580
 QY 882 IleLeuThrSerAsnAsnHisThrGlnProGluHisGlnThrThrLeuThrThrGlu 901
 Db 2581 ATTCTCTACAGCAACACACACACCCAGCCGACACACACAGTGGACACTGCTCACCACAG 2640
 QY 902 GlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgThrPhePheLys 921
 Db 2641 GGAAACATCTTCAGTGCAGAGTCCATGCGCTAGAGAGTGACACTCGGTATTTCTTCAG 2700
 QY 922 MetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIleThr 941
 Db 2701 ATGGAGCCCGCACAGAGTGGGCTGCGCCCTTTCCCGCTGCAGAGTGTGATCTACT 2760
 QY 942 LeuGlnGluThrPheSerAspSerLeuAspValHis 953
 Db 2761 CTGCAAGAGACATCTCAGACTCTCTGGATGTGCAC 2796

RESULT 10
 AB046848

LOCUS AB046848 5549 bp mRNA linear PRI 22-FEB-2001
 DEFINITION Homo sapiens mRNA for KIAA1628 protein, partial cds.
 ACCESSION AB046848
 VERSION AB046848.1 GI:10047332

KEYWORDS Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (sites)
 Nagase, T., Kikuno, R., Nakayama, M., Hirose, M. and Ohara, O.
 Prediction of the coding sequences of unidentified human genes.
 XVIII. The complete sequences of 100 new cDNA clones from brain
 which code for large proteins in vitro
 DNA Res. 7 (4), 273-281 (2000)

JOURNAL MEDLINE 20450683
 PUBMED 10997877
 REFERENCE 2 (bases 1 to 5549)
 Ohara, O., Nagase, T. and Kikuno, R.
 Direct Submission
 Submitted (03-AUG-2000) Oasumi Ohara, Kazusa DNA Research Institute,
 Department of Human Gene Research, 1532-3, Yana, Kisarazu, Chiba
 292-0812, Japan (E-mail: cdnainfo@kazusa.or.jp,
 URL: http://www.kazusa.or.jp/huge, Tel: 81-438-52-3913,
 Fax: 81-438-52-3914)

FEATURES
 source
 1..5549
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="fhl4247"
 /note="vector:pBluescriptII SK plus"
 1..5549
 /gene="KIAA1628"
 <81..3023
 /gene="KIAA1628"
 /note="Start codon is not identified. fhl4247 cDNA clone

for KIAA1628 has a 40-bp different sequence replacing a
 108-bp sequence at the position between 737 and 846, and
 an 88-bp insertion after the position 983 of the sequence
 of KIAA1628, respectively."
 /codon_start=1
 /product="KIAA1628 protein"
 /protein_id="BAB13454.1"
 /db_xref="GI:10047332"
 /translation="MGRKGRSALSRSLRNVQGVKVIIGRGAREGGGQAGTGGG
 GRDLSCGIPRASAPLAAPAIQAPALSRTRASTARFVCRASGRPRPALWHLNG
 APLRNGRVKVGQSGSLVITQIGLDAGYQCVAENSAGMACAAASLAVVVRGLPS
 APTVATPSSSAVLVAMRPEHSSQIIGFSLHYQKARGMDVETQFVAVNNDTTEL
 QVRDLPTNTYEFYVAYISQIAGSRSTPALVHTLDDVPSAAPQLSSSNPSPSIRVA
 WLPLPSSLSQVQVKKYKIEYGLAGKEDQIFSTEVAGNETQLMLNSLPKVKVVRISAG
 TAAGFGAPSMHRTSPMINQSHVPAEALVKQAKMSLVVSGPPPHPTQISGVK
 LYREVCASBEANGDLPLGCGRDQADVPVRLAKVKQVELTOLVGRLYEVKLAVP
 KNGEDGYAAVWKGKTEKAPADMPIQRGPPPLPAHVHAEENSSSTIWLKWKPPFTTV
 NIKEDYVRFPMGLNLSLVITTSGRDLIGLKEPTTKTEFAVQSHGVDMDDPPPS
 VVRSKTLDPSPSTPFDLSPLTFLVRLHMCPTPENGEIVELYLSNHTQPEH
 QMTLLTTQGNIFSAHVHGLSDTRYFFMGARTVEGPFSLQDVITLQKLSDSLDP
 MHSVTGIIVGCLGLCLLACMCAGLARSRRBSLPLGLSSTATPGNALYSRRLGPP
 SPPAAHLSLVHPHDPQWSPDSVDVDRARVHSLMGVSGBSRSHKSKRI SWAQPSS
 LSWAGSWAGCELPQAGPRPALTRALLPACTGTGTLQLLQALVYDAIKNGRKKSPACR
 NOVEARVIVHSDFSASNGNFDLHLEPDPPEAPDLISGDDPQGGAAWLDREL
 GGCLEAAPGPDRLTCLPRAASASCYPDLQPGVLESTPGDSQCLKSPCLGASPLP
 RSPVSSSA"

BASE COUNT 1157 a 1546 c 1496 g 1350 t

Alignment Scores:

Pred. No.: 6,04e-169 Length: 5549
 Score: 4361.50 Matches: 822
 Percent Similarity: 92.55% Conservative: 35
 Best Local Similarity: 88.77% Mismatches: 68
 Query Match: 66.04% Indels: 1
 DB: 9 Gaps: 1

US-09-754-997A-2 (1-1252) x AB046848 (1-5549)

QY 324 AlaAlaAlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeu 343
 Db 243 TCAGCTGCGCCCTCGCGCCCTAGCGGCTCCGCCCATCTCAGGCGCCGCGCGCTG 302
 QY 344 SerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgPro 363
 Db 303 TCGCGGACGGCGGACGACGACGCGCTTCGTGTCCGCGCGCTCGGGGAGCCGCGGCA 362
 QY 364 AlaLeuHisTrpLeuHisAspGlyIleProLeuArgProAlaGlyArgValLysValGln 383
 Db 363 GCGCTGCGCTGCGCTGCACACGCGGCGCGCTGCGGCCCAACGCGGCGCTCAAGTCCAG 422
 QY 384 GlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrTrpGln 403
 Db 423 GCGCGCGGTGGCAGCCTGGTCAACACAGATCGGCGCTGCAGGACGCGCGCTACTACAG 482
 QY 404 CysValAlaGluAenSerAlaGlyThrAlaCysValAlaAlaProLeuAlaValVal 423
 Db 483 TGCCTGCTGAGAACACGCGGGAATGCGTGCCTGCGCGCTGCGCGCTGCGCGTGTGTG 542
 QY 424 ArgGluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerSer 443
 Db 543 GCGGAGGGCTGCCAGCGCCCGCCAGCGGGTCACTGCTACGCCCATGAGCAGCTCCGCT 602
 QY 444 ValLeuValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeu 463
 Db 603 GTGTGTGTGCTGGGAGCGGCCGAGATGCACAGCAGCAGATCATCTCTCTCTCTCTC 662
 QY 464 HistTrpGlnLysAlaArgGlyValAspAsnValGluThrGlnPheAlaValAsnAsp 483
 Db 663 CACTACAGAGGCGCGGCGCATGGCAATGTGAATACCATGTTGCAAGTGAACACAGCAC 722
 QY 484 ThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspThrGluPheTyrValVal 503

Db	723	ACCACAGAACTACAGGTTCCGGACCTGGAAACCAACACAGATTATGAGTTCTACGTGGTG	782
Qy	504	AlaTyrSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThrIleuAsp	523
Db	783	GCCTACTCCAGCTGGGAGCCAGCGCACCTCCACCCACGACTGGTGACACACTGGAT	842
Qy	524	AspValProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAspIleArg	543
Db	843	GATGTCCCAAGTCAGACACCCAGCTCTCCCTGTCCAGCCCAACCTTCGGACATCAGG	902
Qy	544	ValAlaTyrLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLysTyrLysIle	563
Db	903	GTGGCGTGGTGGCCCTGCCCCAGCTGAGCAATGGGAGGTGGTGAAGTACAAGATA	962
Qy	564	GluTyrGlyLeuGlyValAspGlnValPheSerThrGluValProGlyAsnGluThr	583
Db	963	GAATACGGTTTGGAAAGGAAGATCAGATTTCTCTACTAGGTGGGAGGAATAGACA	1022
Qy	584	GlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgIleSerAlaGly	603
Db	1023	CAGCTTATGCTGAACCTGCTTCAGCCAAACAAGGTGTATCGAGTACGGATTTCGGCTGGT	1082
Qy	604	ThrGlyAlaGlyTyrGlyValProSerGlnTyrMetGlnHisArgThrProGlyValHis	623
Db	1083	ACAGCAGCCGGCTCGGGGGCCCTCCCTCAGTGGATGCATCACAGGACGCCCAATGCAC	1142
Qy	624	AsnGlnSerHisValProPheAlaProAlaGluLeuLysValArgAlaLysMetGluSer	643
Db	1143	AACAGAGCCATGTCCCTTTTGGCCCTGCGAGATTGAAGTGGAGCAAGATGGAGTCC	1202
Qy	644	LeuValValSerTyrGlnProProHisProThrGlnIleSerGlyTyrLysLeuTyr	663
Db	1203	CTGGTCTGTATGGCAGCCACACCCCTCACCCACCCAGATCTCTGGCTACAACTATAT	1262
Qy	664	TyrGlyGluValGlyThrGluGluGluAlaAspGlyAspArgProProGlyGlyArgGly	683
Db	1263	TGGCGGAGGTGGGGCTGAGGAGGAGGCCAATGGCGATCGCTGCCAGGGGGCCGTGGA	1322
Qy	684	AspGlnAlaTyrAspValGlyProValArgLeuLysValLysValLysGlnTyrGluLeu	703
Db	1323	GACGAGCTTGGGATGGGGCTGTCCGGCTCAAGAAGAAAGTGAAGCAGTATGAGCTG	1382
Qy	704	ThrGlnLeuValProGlyArgProTyrGluValLysLeuValAlaPheAsnLysHisGlu	723
Db	1383	ACCAGCTAGTCCCTGGCCGGCTGTACAGGTGAGCTCTGGCTTCAACAAACATGAG	1442
Qy	724	AspGlyTyrAlaAlaValTyrLysGlyLysThrGluLysAlaProThrProAspLeuPro	743
Db	1443	GATGGCTATGCAGCAGTGTGGAAAGGCAAGACGAGAGGCGCGCCGACACAGATGCCT	1502
Qy	744	IleGlnArgGlyProProLeuProProAlaHisValHisAlaGluSerAsnSerThr	763
Db	1503	ATCCAGAGGGACACCCCTGCGCTCCAGCCACGTCCTCCGGAATCAACACAGCTCCACA	1562
Qy	764	SerIleTyrLeuArgTyrLysProAspPheThrThrValLysIleValAsnTyrThr	783
Db	1563	TCCATCTGGCTTCGGTGGAAAGACCCAGATTTCCACACAGTCAAGATTGTCACTACCT	1622
Qy	784	ValArgPheGlyProTyrGlyLeuArgAsnAlaSerLeuValThrTyrTyrThrSerSer	803
Db	1623	GTGGCTTACAGCCCTCGGGGCTCAGGAATGCCTCCCTGGTCCCTATTATACACAGTTCT	1682
Qy	804	GlyGluAspIleLeuIleGlyLysValLysProPheThrLysTyrGluPheAlaValGln	823
Db	1683	GGAGAAGACATCTCATTTGGCGGCTTGAAGCCATTCAACAAATACAGATTTCAGTGCAG	1742
Qy	824	SerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSerThrLeuPro	843
Db	1743	TCTCACGGCTGGACATGATGGGCTTTGGCTCTGTGGTGGAGCGCTCCACCTGCCT	1802
Qy	844	AspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrProSerThrVal	863
Db	1803	GACCGGCTTCCACACCCCATCCGACCTGCGACTGAGCCCTGACACCGCTCCACGGTT	1862
Qy	864	ArgLeuHisTyrCysProProThrGluProAsnGlyGluIleValGluTyrLeuIleLeu	883
Db	1863	CGGCTGCACCTGTGTGCCCCCAAGAGCCCAAGGGAGATCTGTGAGTATCTGATCTTG	1922
Qy	884	TyrSerAsnAsnHisThrGlnProGluHisGlnTyrThrLeuLeuThrThrGluGlyAsn	903
Db	1923	TACAGCAGCAACCAACAGCAGCTGAGCACCAGTGGACCTTGTCTCACCGCAGGGAAC	1982
Qy	904	IlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePheLysMetGly	923
Db	1983	ATCTTCAGTGTGAGGTCCATGCGCTGGAGACGACACTCGGTACTTCTTCAAGATGGGG	2042
Qy	924	AlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIleThrLeuGln	943
Db	2043	GGCGGCAAGAGGTGGGACCTGGGCTTCTCCCGCTGCGAGATGTGATCAGCTCCAG	2102
Qy	944	GluThrPheSerAspSerLeuAspValHisAlaValThrGlyIleIleValGlyValCys	963
Db	2103	GAGAAAGCTGTACAGCTCGCTGCATGCATCAGTCACGGGCATCATCGTGGTGTCTGC	2162
Qy	964	LeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGlnSerSerHisArg	983
Db	2163	CTGGGCTCTCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2222
Qy	984	GluAlaLeuProGlyLeuSerSerSerGlyThrProGlyAsnProAlaLeuTyrThrArg	1003
Db	2223	GAATCCCTCCAGGCTGTCTCTCCACCGCCACCCCGGGAAATCCCGGCTGTACTCCAG	2282
Qy	1004	AlaArgLeuGlyProProSerValProAlaAlaHisGluLeuGluSerLeuValHisPro	1023
Db	2283	GCTGGCTTGGCCCCCCCCCAGCCCGCTGCCCATGAATTGGAGTCCCTTGTGCACCCC	2342
Qy	1024	ArgProGlnAspTyrProProProSerAspValGluAspLysAlaGluValHisSer	1043
Db	2343	CATCCCCAGACTGTGTCCCGCCACCCCTCAGACGTGGAGCAGAGGCTGAAGTGCACGC	2402
Qy	1044	LeuMetGlyGlySerValSerAspCysArgGlyHisSerLysArgLysIleSerTyrAla	1063
Db	2403	CTTATGGTGGCGGTGTCTGAGGCGCGAGTCACTCCAAAGAAAGATCTCTCTGGGCT	2462
Qy	1064	GlnAlaGlyGlyProAsnTyrAlaGlySerTyrAlaGlyCysGluLeuProGlnGlySer	1083
Db	2463	CAACCAAGCGGCTGAGCTGGCTGTCTTCTGGCAGGCTGTGAGCTGCCCCAG---GCA	2519
Qy	1084	GlyProArgProAlaLeuThrArgAlaLeuLeuProProAlaGlyThrGlyGlnThrLeu	1103
Db	2520	GGCCCCCGCGCTCTGACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2579
Qy	1104	LeuLeuGlnAlaLeuValTyrAspGlyIleLysSerAsnGlyArgLysLysProSerPro	1123
Db	2580	TTGCTGCAGGCTCTGGTGTACGACGCATTAAGGGCAATGGGAGGAGAGTCAACCCCA	2639
Qy	1124	AlaCysArgAsnGlnValGluAlaGluValIleValHisSerAspPheGlyAlaSerLys	1143
Db	2640	GCCTGCAGGAACCCAGGTGGAGCTGAAGTCAATGTCCACTCTGACTTTAGTGCATCTAAC	2699
Qy	1144	GlyCysProAspLysHisLeuGlnAspLeuGluProGluLeuProLeuThrAlaGluThr	1163
Db	2700	GGGAACCTTGACTCTCATCTCAAGACCTGGAGCCCTGAGGACCCCTGCTCCAGAGGCT	2759
Qy	1164	LeuProSerThrSerGlyAlaValAspLeuSerGlnGlyAlaAspTyrLeuGlyArgGlu	1183
Db	2760	CCTGATCTCATCTCGGCTGTGGGATCCAGGGCAGGGGCGGCTGCTGGTGGACAGGGAG	2819
Qy	1184	LeuGlyGlyCysGlnProThrThrSerGlyProGluArgLeuThrCysLeuProGluAla	1203
Db	2820	TTGGAGGGTGTGAGCTGGCAGCCCCCGGGCCAGACAGACTTACCTGCTTGTCCAGAGCA	2879
Qy	1204	AlaSerAlaSerCysSerCysSerAspLeuGlnProSerThrAlaIleGluAlaPro	1223
Db	2880	GCCAGTGTCTCTGCTCTTACCCGGACCTCCAGCCAGGGGAGGTGCTAGAGGAGACCCCT	2939

Qy 1224 GlyysSerCysGlnProLysAlaLeuCysProLeuThrValSerProSerLeuProArg 1243
 Db 2940 GGAGTAGCTCCAGCTCCTATCCCTCGCTCTAGGAGCCAGCCAGGCTGCCAGCA 2999
 Qy 1244 AlaProValSerAla 1249
 Db 3000 TCCCGGCTCTCTCTCT 3017

RESULT 11
 AC110235/c
 LOCUS
 DEFINITION Mus musculus clone RP23-100M12, WORKING DRAFT SEQUENCE, 17 ordered pieces.
 ACCESSION AC110235
 VERSION AC110235.5 GI:20800205
 KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Mus musculus, clone RP23-100M12
 Unpublished
 2 (bases 1 to 158055)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kellis,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McSwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

DIRECT SUBMISSION
 Submitted (10-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 158055)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kellis,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McSwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS

TITLE
 JOURNAL
 COMMENT

Direct Submission
 Submitted (15-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On May 15, 2002 this sequence version replaced gi:20531900.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L22383
 Center clone name: 100 M.12
 ----- Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 153807 bases at least Q40
 Consensus quality: 155134 bases at least Q30
 Consensus quality: 155784 bases at least Q20
 Insert size: 218000; agarose-fp
 Insert size: 156455; sum-of-contigs
 Quality coverage: 7.1 in Q20 bases; agarose-fp
 Quality coverage: 10.0 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 17 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

This sequence will be replaced
 by the finished sequence as soon as it is available and
 the accession number will be preserved.

1 6459: contig of 6459 bp in length
 6460 6559: gap of 100 bp
 6560 6667: contig of 108 bp in length
 6668 6767: gap of 100 bp
 6768 7863: contig of 1096 bp in length
 7864 7963: gap of 100 bp
 7964 9426: contig of 1463 bp in length
 9427 9526: gap of 100 bp
 9527 11011: contig of 1485 bp in length
 11012 11111: gap of 100 bp
 11112 12741: contig of 1630 bp in length
 12742 12841: gap of 100 bp
 12842 14789: contig of 1948 bp in length
 14790 14889: gap of 100 bp
 14890 17826: contig of 2937 bp in length
 17827 17926: gap of 100 bp
 17927 21244: contig of 3318 bp in length
 21245 21345: gap of 100 bp
 21346 27413: contig of 6069 bp in length
 27414 27513: gap of 100 bp
 27514 36381: contig of 8868 bp in length
 36382 36481: gap of 100 bp
 36482 45104: contig of 8623 bp in length
 45105 45204: gap of 100 bp
 45205 58369: contig of 13165 bp in length
 58370 58469: gap of 100 bp
 58470 85782: contig of 27313 bp in length
 85783 85882: gap of 100 bp
 85883 112706: contig of 26824 bp in length
 112707 112807: gap of 100 bp
 112808 153927: contig of 41121 bp in length
 153928 154027: gap of 100 bp
 154028 158055: contig of 4028 bp in length.

Location/Qualifiers
 1. 158055
 /organism="Mus musculus"
 /mol_type="genomic DNA"

FEATURES
 source

/db_xref="taxon:10090"
/clone="RP23-100M12"
/clone_lib="RPCI-23 Female Mouse BAC"

misc_feature

1..6459
/note="assembly_fragment
clone_end:596
vector_side:left"

misc_feature

6560..6667
/note="assembly_fragment"

misc_feature

6768..7863
/note="assembly_fragment"

misc_feature

7964..9426
/note="assembly_fragment"

misc_feature

9527..11011
/note="assembly_fragment"

misc_feature

11112..12741
/note="assembly_fragment"

misc_feature

12842..14789
/note="assembly_fragment"

misc_feature

14890..17826
/note="assembly_fragment"

misc_feature

17927..21244
/note="assembly_fragment"

misc_feature

21345..27413
/note="assembly_fragment"

misc_feature

27514..36381
/note="assembly_fragment"

misc_feature

36482..45104
/note="assembly_fragment"

misc_feature

45205..58369
/note="assembly_fragment"

misc_feature

58470..85782
/note="assembly_fragment"

misc_feature

85883..112706
/note="assembly_fragment"

misc_feature

112807..153927
/note="assembly_fragment"

misc_feature

154028..158055
/note="assembly_fragment
clone_end:17
vector_side:right"

BASE COUNT 41469 a 36830 c 36521 g 41635 t 1600 others
ORIGIN

Alignment Scores:

Pred. No.: 5,71e-117 Length: 158055
Score: 3125.50 Matches: 913
Percent Similarity: 23.34% Conservative: 2
Best Local Similarity: 23.29% Mismatches: 6
Query Match: 47.33% Indels: 3002
DB: 2 Gaps: 13

US-09-754-997A-2 (1-1252) x AC110235 (1-158055)

QY 332 AlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAla 351
|||||
DB 85019 GCTGCCCCAGCCATCTCGCAGGCGCCCGAGGCGCTCTCGCGACGCGGGCCAGCAGCGG 84960
QY 352 ArgPheValCysArgAlaSerGlyGluProArgProAlaLeuHisTrpLeuHisAspGly 371
DB 84959 CGCTTCGTGCGCGGCGTCCGGGAGCCACGCGCGCGTGCCTGGCTGCACGACGCG 84900
QY 372 IleProLeuArgProAnGlyArgValLysValGlnGlyGlyGlySerLeuValIle 391
DB 84899 ATCCCGTTCGACCCCAATGGCGCGCTCAAGGTGACGGCGGTGGCGGACGCTTGGTCATC 84840
QY 392 ThrGlnIleGlyLeuGlnAspAlaGlyTyrTyrGlnCysValAlaGluLeuSerAlaGly 411
DB 84839 ACTCAGATCGCCCTTCAGGACGCTGGCTACTACAGTGCCTAGCAGGAGGAGGAGGAGG 84780
QY 412 ThrAlaCysAlaAlaAlaProLeuAlaValValArgGluGlyLeuProSerAlaPro 431
DB 84779 ACTGCCTGTGCGCGCCCTGGCGGTAGTGGTGCAGGCGGCTGCCACGCGCCCG 84720

QY 432 ThrArgValThrAlaThrProLeuSerSerSerValLeuValAlaTrpGluArgPro 451
DB 84719 ACTCGGCTCACAGCACCGCGCTGAGCAGCTCTCTGTGTGTGTGGCTGGAGCGGCT 84660
QY 452 GluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTyrGlnLysAlaArgGlyVal 471
DB 84659 GAGTTGCACAGCAGCAAAATCAATTGGCTTCTCTCTCTACACCAAAAGCAAGGGGTAT- 84601
QY 471 ----- 471
DB 84600 GTCTACTCTTGACCGCGGAATCCAGTTAGGGCAAGTATGGGATGGATTAATTAAGATG 84541
QY 471 ----- 471
DB 84540 AGAGGAAGGATGGGAGCAATACTGCTGGGAAGTATCTAGTATTAGAGAATGGAGTGT 84481
QY 471 ----- 471
DB 84480 ACACGCACTCCTTGACGGAGAGAAATGCTTCTCTTTCAGGGTTAAAGGAGCCATCTTCCA 84421
QY 471 ----- 471
DB 84420 CCCCCAGCTAACTAACTGCAGGACGGGCTCCTTCAATTAGCAGGGGAGCTTCTGGGCT 84361
QY 471 ----- 471
DB 84360 ATTATGGAATCAATGCATTAAGTAGACTAGATGGCAGAGGCAATTTCTATCTTCCACCCTA 84301
QY 471 ----- 471
DB 84300 GATTGCTGTAACTTTAGCTATACATAGACCAGTAGGATCCCGACAGCTCTCCAGCT 84241
QY 471 ----- 471
DB 84240 AGTCAGACATGAGTGGGCTGTGACAGTCCCTTTGTCCCCCGTGTACACCTCAGCA 84181
QY 472 --- AspAsnValGluTyrGlnPheAlaValAsnAsnAspThrThrGluLeuGlnValArg 490
DB 84180 GTGGACATGTGGAGTACCAGTTTGCAGTAAACATGACACACAGAGCTGCAGGTTCGG 84121
QY 491 AspLeuGluProAsnThrAspTyrGluPheTyrValValAlaTyrSerGlnLeuGlyAla 510
DB 84120 GACCTGGAACCCACACCGAATTATGAGTTCTACGTGTGTGGCTACTCCAGCTGGGGCC 84061
QY 511 SerArgThrSerSerProAlaLeuValHisThrLeuAspAsp----- 524
DB 84060 AGCCGAACCTCCAGCCCGAGCTGTGTGCATACACTGGACGATGGTAGGACCTCTGAACCT 84001
QY 524 ----- 524
DB 84000 GCAGTGACCTGGGACCCAGAGACACTCCCACTCAGGGACAGTGTGGCCCAAGGCCCTCTT 83941
QY 524 ----- 524
DB 83940 CCCTGCCTTGTGATGTCTGTGTGCTTATAAAGCAGTGGGCGAGAGTCTTCTTCCAG 83881
QY 524 ----- 524
DB 83880 TGTAAATGGGAAACTGGGGCTCAGAAATGAAATGACTTTGCCAAGCTCAGCAGGCA 83821
QY 524 ----- 524
DB 83820 GCTGGGACTTGAGCCAGCTCTGAGCCCCCATACCATGCGAGCCCTTCTTCTTAACATACA 83761
QY 524 ----- 524
DB 83760 TTACTACCCAGCCATTGGATATGAAAGCTTCTGGCCACACATGTCCACTAGCCTGGGA 83701
QY 524 ----- 524
DB 83700 TAGAAAGGATGTCTCGCATCTCCAGCGTAGGGCTCCCTGAGAGGCTCCCTCCCTCC 83641

Qy	524	-----	524
Db	83640	TCCTCAGATCTGGGCTGGGCTACTTCTGACTGCTGTGACCTGGGCAAGTCTCT	83581
Qy	524	-----	524
Db	83580	GATCACTCCGACTGTGCTTATCTTCCTCTGGTGGAGGGAGCTCAGCCAAATGTTTC	83521
Qy	525	-----	541
Db	83520	TTCTGTCTCAGTCCCGAGCGCAGCACCCAGCTTACCTTGTCCAGCCCCCAACCCCTCGGA	83461
Qy	541	pIleArgValAlaTrpLeuProSerSerLeuSerSerLeuSerAsnGlyGlnValLeuIleuIysTy	561
Db	83460	CATCAGGGTGGCATGGTGGCTCCCTGCTCCCTCCAGCTCAGCAATGGACAGGTGCTGAAGTA	83401
Qy	561	rIyAlIeGluTyGlyLeuGlyIys	569
Db	83400	CAAGATAGATACGGTTTGGGGAAGGAAGTGAAGTGGGCGCGGGAGGTATACAGAGATGTC	83341
Qy	569	-----	569
Db	83340	AGTGGGAATGGGGTCTCCCTAGATGAGGAGCGTTGTCAGTGGAAGTGGAGCCTAATT	83281
Qy	569	-----	569
Db	83280	GGGGCATGGATGAGAAAGCTCTGGAAGAGCAATTGGGAGACTAGATAGCCCAATGTGT	83221
Qy	569	-----	569
Db	83220	AAAGTATGTGTGCAAGCCCAAGGCTGAGTTCAATCTCAGACTTCATGTTAAAAA	83161
Qy	569	-----	569
Db	83160	AAAGCTGGCACAGTGTGAGTGCTGCTGTGTAATCCAGAGCAGGGAGGTGGAGACAC	83101
Qy	569	-----	569
Db	83100	GGGGACACCTTGAGCTCACTGGGGTAAACCAAGTGAAGTTCCAGGCCAGGAGTCCCTGT	83041
Qy	569	-----	569
Db	83040	GCAGAAGAGAGGTGGCAATACCAAGAGAAATGACAAACAGAGGTTGTCTGACCTCCA	82981
Qy	569	-----	569
Db	82980	CATGTATGTGTACACACCTGCACATAGTGGAACACTCACATCTATGGACACACAGCGA	82921
Qy	569	-----	569
Db	82920	CACAGGGTTTGGGGCAGTTGAGGAAAGGGAAGCAAGAGGGTCAGTAGCTGGCCAC	82861
Qy	569	-----	569
Db	82860	AGATAAGGAGGTGATCAGTGGGACCTTGGACTGTTCCCTTCTGTGAAGGAGCAATTT	82801
Qy	569	-----	569
Db	82800	ACCAGTGGATCCTTGGAAACCTTCTAGGTTTTTGTGTTTTTAAAGATGTAT	82741
Qy	569	-----	569
Db	82740	TTATTTATTTATATATAGATTCTCTATCTGACGCCCACTGTCATGCCAAGAGAGGG	82681
Qy	569	-----	569
Db	82680	AATCAGATCCAGTACAGATGGTTGTGAGCCACAGTGTGTGCTGGGAATTTGAATCTCAG	82621
Qy	569	-----	569
Db	82620	GACCTCTGGAAGACAGTGTGCTCTTAACAATTTGAGCCATCTCTCCAGCCCCAGAAAC	82561
Qy	569	-----	569

Db	82560	CTTCTGGTTTTACACACACAGAGAAGATGTGGAGGCTGCAGGAGTGACATGACATGGG	82501
Qy	569	-----	569
Db	82500	CTGCTCCTACTGGGAGATGATGCTGCTCTGGGGAGGTGATGTGTGGAAAGGAATG	82441
Qy	569	-----	569
Db	82440	GAGAGTGAGGAAAGAAAGTGTGTAGGGTCAAGCTCAGGGTAAAGCTAAAGAGAGAGAGCAGGAGG	82381
Qy	569	-----	569
Db	82380	CTGATGGAGACAGCAGTCAAGGCTCTCTGACACAGGGACTGCTTAGTGTCTTATGCCAG	82321
Qy	569	-----	569
Db	82320	CCATCAAGCCTGTGTAAAGAGGTGTGTACTGGCTGTGTGCTGTTCACAGGGAGGAA	82261
Qy	569	-----	569
Db	82260	GTTTCATGCTTGTTCCTCATATGGAACGAGGCCAATGGGCTTCAGCAGCGATAGGTA	82201
Qy	569	-----	569
Db	82200	GAAGCCAGGGTCAACCTGGAGGCATGAACCTGCTCCGTGTGAGGGGTCAGTGCCCA	82141
Qy	569	-----	569
Db	82140	GAGGACCTCGGGACCTGAATGGAAGTCAATGTGATAAATTTTAAAGATTGATGTCC	82081
Qy	570	-----	576
Db	82080	CCCAATGAAGCACCCAGCAGCCTATCTTTTGTCCCAAGAGATCAGGTTTCTCCAC	82021
Qy	576	rGluValProGlyAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnIysValTy	596
Db	82020	CGAGGTGCTGGAAATGAGACACAACTTACGTTAAACTCACTTCAGCCAAACAAAGTGA	81961
Qy	596	rArgValArgIleSerAlaGlyThrGlyAlaGlyTyGlyValProSerGlnTrpMetGl	616
Db	81960	CCAGTCCCGATTTCACTGAGCTGGCACTGGCGCTGCTATGGAGTCCCTTCTCAGTGGATGA	81901
Qy	616	nHleAsnThrProGlyValHleAsnGlnSerHle	627
Db	81900	GCACAGACACCTGGTGTGCAACAACCAAGGCCATGGTATGGACCTACACGGGCAGACTGG	81841
Qy	627	-----	627
Db	81840	ACAGCGGGACAGGTGCTCTGTGGCTTAGATATTGCCCTGTATTGCTGTTATTGCC	81781
Qy	627	-----	627
Db	81780	ACCCCCCCCCCCCCACACACACTTGTATGACCCCGAAGGCTCGGTGTAGGAGTGA	81721
Qy	627	-----	627
Db	81720	AGGGCTGGGCTCAAAATATCTCTGTGTCACTCTGTGACTTTTGACAAGCCACTCGCATCTC	81661
Qy	627	-----	627
Db	81660	CAGAGCTTCGGTACTACCAAGGATGATAGCAGGACTAGTCACTTGTAAATTATACACAGAT	81601
Qy	627	-----	627
Db	81600	GGTCCGATAGAACAGGAGCTCTTGGGTTGAGACCAAGGCTCACTTGCAGAGTGTCTG	81541
Qy	627	-----	627
Db	81540	CCTAGCATGTACAAAGCACTAAATTCATCCCGAGGACCAATAAACCCAGCATACCTCT	81481
Qy	627	-----	627

QY	740	-----	740	
Db	79260	GTATCTGCTCTCTGGAGCTTATTTATGCTTATCTAATAACATTTGTTCATCACCAAT	79201	
QY	740	-----	740	
Db	79200	AGTATACACTGTCCACAGCTCTTACCTGCTTTGTAAGCAGCCTCTTAAGGACACTCAT	79141	
QY	740	-----	740	
Db	79140	GAGGGGACAGTGAGATGGTTTCAGCAGGGAAGGTGATTGCTTACAAGCCACATGACCTGAGC	79081	
QY	740	-----	740	
Db	79080	TTGGTTTTCGACGCCCCCCCAAAATGTCTCTGCTCCCCCACAATGCACACAGCAGCACA	79021	
QY	740	-----	740	
Db	79020	CTCCCTCCCTCCCTCCACCAATTAATGTGTGACTGAGAAGATGCAGGAGTTAAGAAAT	78961	
QY	740	-----	740	
Db	78960	TTAAATATGTTTAAAGATGATTTATGATTATATGTTGAGTACACTGTCTCACTGTTG	78901	
QY	740	-----	740	
Db	78900	CTATCTTCAGACACAGCAGAGAAGGTCACTAGATCCCAATTATAGATGTTGTGAGCCACC	78841	
QY	740	-----	740	
Db	78840	ATGTGTTGCTGGGAATTGAATCTCAGGGAATCTGGAAGACAGTCACTGTTAAACCGCTG	78781	
QY	740	-----	740	
Db	78780	AGCCCTCTCTTCAGCCCCCAGATTTTAAATATTTTAAAGAGTAAATATAAGAAAGA	78721	
QY	740	-----	740	
Db	78720	AGAAATGTCAAGATGTAGGTATTTGTCATCAGATCCCTTATAAGAAATAGAAATGATGGC	78661	
QY	740	-----	740	
Db	78660	ACAGAGAGGTTAAACAACTTCTCTAAAGCCACATACATAGCTAGGAAGTGGCAGCAATGG	78601	
QY	740	-----	740	
Db	78600	AGCTGGCTCCAGAGTCAGGGCATGCTTCTCTGTGCATCTCTCATCAGCCATGGGATAT	78541	
QY	740	-----	740	
Db	78540	GTACAGTCACTCTATGTGTGACCTAAGCAAAAAAATAATCACTTCCATTCTTGGGAGCCT	78481	
QY	740	-----	740	
Db	78480	CAGTCGAGTAGGAATATGTACCCCAACTCAAGGATCAAGAGAGGCCCGGTGCCAG	78421	
QY	740	-----	740	
Db	78420	CTGTGTTATCACACAGCTGTAAACCCCTGTCTAAAGACCAGAGGCACAGGATGGAA	78361	
QY	740	-----	740	
Db	78360	TTACAGCCAGCCTTAGGCTACTACATAATGAGATTCTGCCTCAAAATAAGAGGGGAAA	78301	
QY	740	-----	740	
Db	78300	GGCCAAATGAACCTACTAGACACCATGATGCTCCACCTACTCTCTGCACACTTGGGTTTC	78241	
QY	741	-----	758	
Db	78240	CCCTCTAGACCTTGCCTATCCAGAGGGGCCACCGTGCCTCTCTGCCATGTCCAGCGAG	78181	
QY	758	luSerAenSerSerThrSerIleTrpLeuArgTrpLysProAspPheThrVal	778	
Db	78180	AGTCAACAGCTCCCACTTCATTTGGCTTCGGTGAAGAGCCAGACTTTTACCACTGTCA	78121	
QY	778	YelleValAenTrpThrValArgPheGlyProTrpGlyLeuArgAenAlaSerLeuVal	798	
Db	78120	AGATTGTCAACTACACTGTACGCTTCGGGCCCTTGGGGGCTCAGGAATGCTTCCCTGGTCA	78061	
QY	798	hrTrpTrpThr	801	
Db	78060	CCTACTATACAGGATATCAGTGAGGGGGAGTGTAGCTGGGGTGGTAGACTAGCGGGTTC	78001	
QY	801	-----	801	
Db	78000	AGCTGTGTGAGGCCACACCCAGAGATCTACACAGATGTCTTTTCGTGACTGCTGACCTG	77941	
QY	801	-----	801	
Db	77940	GGGGTAAATAGAGGATGGGATTTCCCCCACTCCCTCAGTGGCAGAAATGTGTGGTCAGCC	77881	
QY	801	-----	801	
Db	77880	TTTGTGAGAGTGGATGCAAGACTAGAGGCTTCCATCCGTTCGTGGCACCCCTCCCGGCC	77821	
QY	801	-----	801	
Db	77820	CATCATTCGGAAGATAGCTTAAGAGAGCTGCTGTCTGTGATTGGGAAGACACCCCTGTGAAG	77761	
QY	801	-----	801	
Db	77760	CCAAACAGAGAGAAACTAAGCTGAGGATAGACACTTCAGGAGCGATTTGTTCCCCAGGT	77701	
QY	801	-----	801	
Db	77700	GCAGAATACTTAGATTAATTCGGGAAAGCGTCCAGAGAGAGAGAGGCTTGTCTTTGGA	77641	
QY	801	-----	801	
Db	77640	GATGAGACATGGAATAACATATGTCCGGCCTTTCTGTGGTTCCTGCTGTGTGTGGGA	77581	
QY	801	-----	801	
Db	77580	GGGAGGTTGTGGGCCCAACAGACAGCTCCCTCCAGGTTGGGGCTGGAGGAGGG	77521	
QY	801	-----	801	
Db	77520	GCTTCTGTCTGGGTTCTGCTGGCTTGAGGCTGTCTCGAGGCACTAGTCTTATCCCTT	77461	
QY	802	-----	819	
Db	77460	CTCCTGGAAGCTCTGGAGAAGACATTTCTCATTTGGCGGCTCGAAACCATTTTACCAAGTACG	77401	
QY	819	luPheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGlu	839	
Db	77400	AGTTTGGGTACAGTCCCAAGAGTGGATATGATGGGGCTTTTGGCTCCGTGTAGAAC	77341	
QY	839	rgSerThrLeuProAspArg	845	
Db	77340	GCTCCACCTGCTGTGACCG-TGATATCCACTGCCCGCTCCACCTGTGCCACCATTTG	77282	
QY	845	-----	845	
Db	77281	CTGGGTCTGAGGAAGAGAGAGAGGAGGCTGTGCCCTCATCTCCCTTTAGTAGGCTG	77222	
QY	846	-----	853	
Db	77221	TGTCTGTACTCTCCACCCCTTCCCTCTGTGTGAGGGGCTTCAACACTCTCTCTGACC	77162	
QY	853	euArgLeuSerProLeuThrProSerThrValArgLeuHisTrpCysProProThrGlu	873	
Db	77161	TGGCCTGAGCCCTTGACACCATCCACGTTCCGTTACCTGTGTCCCTCCACCGAGC	77102	
QY	873	roAenGlyGluIleValGluTrpLeuIleLeuTrpSerAenHisThrGlnProGlu	893	

Db	77101	CCAAATGGTGAGATTGTGGAGTATCTAAATTCTCTACAGCAACAACACACCCAGCCGCAAC	77042
Qy	893	ieGlnTrpThrLeuLeuThrThrGlu-----	901
Db	77041	ACCAAGTGAACACTGCTCACCACAGA- GGGAAGGAGGCCACCCCTGCTCTCTGCCCCTT	76983
Qy	901	-----	901
Db	76982	TTGTTCCAAAGGTCTCTGGGATTGGAGGCTGAAGCTACTGTCTGGGTGCAAAAGTTGGTTT	76923
Qy	902	-----GlyAsnI 904	
Db	76922	GCCCAATGGTCTTGGCCCTCTGCGCACCCCTGCTCCTCCCACTTCCCAACCAAGAAACA	76863
Qy	904	lePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePheLeuMetGlyA 924	
Db	76862	TCTTCAGTGCAGAGTCAATGGCCCTAGAGGTGACACTCGGTATTTCTTCAGATGGGAG	76803
Qy	924	laArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIleThrLeuGlnG 944	
Db	76802	CCCGCACAGAGGTGGGGCTTGGGCCCTTTTCCCGCTTGCAGGATGTGATTACTCTGCAAA	76743
Qy	944	luThrPheSer-----	947
Db	76742	AGACATTCTCAGGTACCAAGGCAGAGGAGAGAGGCCCTTGGGACAACCTGGAGGCCCAAG	76683
Qy	947	-----	947
Db	76682	ACCCCTGGATTGGTCCCAAGGCCCTGCTTCTGGAAATTAATTTCTTTACTTTGGAA	76623
Qy	947	-----	947
Db	76622	CCCTAAGTGTGGGGATGGAGAAGCAGTGACAGCTTGATAGAGGACTGTCTATAGCTTT	76563
Qy	947	-----	947
Db	76562	GAGGGTCTTTTATTATTATTCTCTGTGTGTGTTTGGCACAGTGATATGCAAGG	76503
Qy	947	-----	947
Db	76502	TCAGAGGACATCAGAGAGTTGGTTTGTGGGGATCTTGCTTTGCTGGAGGGTTGGGTGA	76443
Qy	947	-----	947
Db	76442	GCTCCTGAGTCTGTGMAAGTGTGCAATAGGCTGGATAGACACTCACAGGGCTCAAT	76383
Qy	947	-----	947
Db	76382	TAAGATACAAATGAGATCCAGGGCTCTTATTGTGTAATTTATGAAGATGTGAAGATGA	76323
Qy	947	-----	947
Db	76322	GCCAGTGGGGCATTTAGCCACACATAGGACTGAGAGGATGAAGCCCTGAGTCCAAATGCA	76263
Qy	947	-----	947
Db	76262	CACCTGAGTGTGCTGTCT	76203
Qy	947	-----	947
Db	76202	GTAATCCTTGTGAGGATGACAGACAGGATATTCTGGGGCTGAGAGGGAAGCTGGGGAAG	76143
Qy	947	-----	947
Db	76142	TCAGGTGATGTAGACCTCTGTCTCCACTACATATACATCCCTTATAGTGAACATATA	76083
Qy	947	-----	947
Db	76082	CAGACATGACACACATGTGTGCACCTAGAGCCCACTTGTATTATATATATACACCA	76023
Qy	947	-----	947
Db	76022	TGAGAGCCCTTGTGGCAGGAGGMAAACAACTTCACTGAAGAGCTGATTCTATAGATCC	75963
Qy	947	-----	947
Db	75962	AGTTTTGTGAGAGTGGCAGGTTCTGTATCAGCAGGGAAGACCGTGTGTGACTCAGTATGA	75903
Qy	947	-----	947
Db	75902	GGCCCAAGTAAAGTCCCGTGGCTCTCTCTATTCCAATCAGAAGGAGTGAGCTTCCAAACC	75843
Qy	947	-----	947
Db	75842	AGGTGTGAACCAAGTGTGGTACTGTGCACACCTTTGTATCTCGTCTCTGTGTTATTGAGCACTG	75783
Qy	948	-----AspSerLeuAspValHisAlaVal 955	
Db	75782	CTTCTCTTGCACCTGGAGTCACACTGTTCTCCCCCGCAGACTCTCTTGATGTGCACCGC	75723
Qy	955	alThrGlyIleIleValGlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysA 975	
Db	75722	TCAGGGCATCATCGTGGGTGTCTGCTGGGCTTCTCTGCTCTCTGCTGCTGCAATGTGTG	75663
Qy	975	laGlyLeuArgGlnSerSer-----	981
Db	75662	CTGCTTACGACGAAGCTCCACAGGTGAACCGGGGAACCACTGAGACCACTGAGTTACC	75603
Qy	981	-----	981
Db	75602	AGGAAGGCTGCAGAGTGGACAAAGAGCGTGTCTTCTCTGCTGCCCCCGGAAGCCCAAG	75543
Qy	981	-----	981
Db	75542	TATAGATAGGGCTATGAGCTGAGTGAAGTTCACTGTGCTACATGTGCCCCCATGTGAGGC	75483
Qy	981	-----	981
Db	75482	CCTGGGCGTCAATCTCTCAGTAGACACAAACAAGTGGATAATTTATTATTAGGATTTCTGAG	75423
Qy	981	-----	981
Db	75422	ATGAAGGATTGAGGTGTGTTCTCTCATGAGAATGTAGAATAAACTTAGGGAAGCCCAAG	75363
Qy	981	-----	981
Db	75362	CTCCCACTCAGTAAACAGCGGGCTGAGTACCTGTGGGCTTCCCTCATGTCTTAAATGG	75303
Qy	981	-----	981
Db	75302	AAATGGGCCATCTCGCCATACAGGAAGCGTGCCATCTGAGGGGTGAGCCGAGTGGACT	75243
Qy	981	-----	981
Db	75242	TGTATTCTCTTCTCTGTTTGTGATTGTCACACACCGAAGAACCAATCCACTCTTGTCCG	75183
Qy	982	-----HisArgGluAlaLeuProGlyLeuSerSerSerGlyT 994	
Db	75182	ACCCATAACTTTGTTCTGTTGATCACAGGAAGCCCTGCCCCGATTTGCTCTCAGGCA	75123
Qy	994	hrProGlyAsnProAlaLeuThrArgAlaArgLeuGlyProSerValProAlaA 1014	
Db	75122	CCCCAGGAACCCAGCGCTCTACACAGAGCTCGACTTTGGGCCCCCAGTGTCTCTGCTG	75063
Qy	1014	laHisGluLeuGluSerLeuValHisProArgProGlnAspTrpSerProProSerA 1034	
Db	75062	CCCATGATTGAGTTCCTCTGTGATCTCTGCTCCCAAGATTGGTCCCCCACCCTCTAG	75003
Qy	1034	spValGluAspLysAlaGluValHisSerLeuMetGlyGlySerValSerAspCysArgG 1054	
Db	75002	ATGTGAAGACAAGCTGAAGTACACAGCTTATGGGTGGCAGTGTGTTTTCAGATTGCCGG	74943
Qy	1054	lyHisSerLysArgLys-----	1059
Db	74942	GCCACTCCAGAGAAGGTGAGACTGGAGTCTCTTTCAGGTGAGGGACTCACACGGAACCCA	74883

QY 1059 ----- 1059
 Db 74882 TTCCAAATCAAAACACGAGGGGACACGAGACACAACCTTCTGAGGTGAGATGCGCTCGC 74823
 QY 1059 ----- 1059
 Db 74822 GCGCAC 74763
 QY 1059 ----- 1059
 Db 74762 ATACAGAGAAAGCCTGTCCATACACCTGGGAGGATCCAGAGCTGTGTTTTCGAGAGCA 74703
 QY 1059 ----- 1059
 Db 74702 CCCATCTTCTTCACTTACCTAATATGGAAGAAAGAGGCTGTCTCACTTTTGTGTGGC 74643
 QY 1059 ----- 1059
 Db 74642 CAAGTAGGGTGGCCCCCAGATAAGGTGAACCCCATAGGCCAGGTGTTCTGGGGGACACA 74583
 QY 1059 ----- 1059
 Db 74582 CCCTGACCAAGTGACAGTCAGTCTTCCCATCTTCTCACTTCTCTGCAATTAGAAGA 74523
 QY 1059 ----- 1059
 Db 74522 GTCCAGCAAGCTTGGGAGCTTTGGATTGTGCACATCAACTGTTGTAGGCTTATGGGTCC 74463
 QY 1059 ----- 1059
 Db 74462 GCCCTTCCCATTTCCCTACTCACTAGTCTCACTCTGCTCCCGGGTGGGGCTTTAGTGTG 74403
 QY 1059 ----- 1059
 Db 74402 GAGTTTAGTGTATCTGGAGATAACAGGGTCCACAGGTCTGGACCTGCGATGCCCATAC 74343
 QY 1060 ----- IleSerTrpAlaGlnAlaGlyGlyProAenTrpAlaGly 1072
 Db 74342 TGTCTTCTTCTCTCTGAGATCTCTCTGGCTCAGGAGGGGACCAAACTGGCGAGGC 74283
 QY 1073 SerTrpAlaGlyCysGluLeuProGlnGlySerGlyProArgProAlaLeuThrArgAla 1092
 Db 74282 TCCTGGGCGAGCTGTGAGCTGTCCCGCCAGGGTAGTGTCCAGAGCGGGCTCTGACCCCGTCT 74223
 QY 1093 LeuLeuProAlaGlyThrGlyGlnThrLeuLeuGlnAlaLeuValTrpAspGly 1112
 Db 74222 CTGCTGCTCCAGCGGAAACGGGACAGCACTGCTGCTGCAAGCCCTGTTATATGACCC 74163
 QY 1113 IleIys ----- 1114
 Db 74162 ATAAAGTGTGTTGTTGCTGTTTGGAGAGGGAGGAGGAGGATGGGAGGCGAGGG 74103
 QY 1114 ----- 1114
 Db 74102 AAAGTCCCCCTCCATGCACTTCCAGGACAGAAAGTGTATTTGAAAGACCTTGTAGCTTC 74043
 QY 1114 ----- 1114
 Db 74042 CAAAGGGAGTGGTGACAGAAAGCAGTGAGGACTTCTTTACCTGGTCTTAAGTGACT 73983
 QY 1114 ----- 1114
 Db 73982 TACAGTAGACTTACAGAGTATCTGAAGGTGGCCTTAGCCCTTAGATACGGCCAGCGCCT 73923
 QY 1114 ----- 1114
 Db 73922 CCTTGTCTCTGGAGGTTCTGGGTCTAGTACTTCTGTGTAGCTGAGTGGGCGCAACG 73863
 QY 1114 ----- 1114
 Db 73862 TGTGACATTTCTGGTTTCTCATCTCTGCCAGACGCTGAGCGTGTCTTCTTCTTAGGTGGA 73803
 QY 1114 ----- 1114

Db 73802 ACAGTGTCTTGTATCATTAATGGGAGTCAAGCTCTCCCCCGCCCCCAATATATTGGT 73743
 QY 1114 ----- 1114
 Db 73742 CATTTTCTGAGACCAATTTCTAGTGTGTCTCACTAGCCTTTTGTCTCACTCTGCCAG 73683
 QY 1115 SerAsnGlyArgIysLysProSerProAlaCysArgAsnGlnValGluAlaGluValIle 1134
 Db 73682 AGCAACGGGAGAAAGACCGCTCCCGAGCTGCGAGGAATCAGGTGGAAGCTGAGGTCAATT 73623
 QY 1135 ValHisSerAspPheGlyAlaSerLysGlyCysProAspLeuHisLeuGlnAspLeuGlu 1154
 Db 73622 GTCCACTCCGACTTCGGTGCATCCAAAGGATGTCTTGACCTCCACCTCCCAAGACCTGGAG 73563
 QY 1155 ProGluGluProLeuThrAlaGluThrLeuProSerThrSerGlyAlaValAspLeuSer 1174
 Db 73562 CCAGAGAACCACTGACTGCAGAGACTTCGCTTCACAGTCTGAGCTGTGGATCTGTCT 73503
 QY 1175 GlnGlyAlaAspTrpLeuGlyArgGluLeuGlyGlyCysGlnProThrThrSerGlyPro 1194
 Db 73502 CAAGGAGCAGACTGGCTGGGCGAGGCTGGGAGGTGCCAACCAACCACTGGGCCA 73443
 QY 1195 GluArgLeuThrCysLeuProGluAlaAlaSerAlaSerCysSerCysSerAspLeuGln 1214
 Db 73442 GAGAGGCTCACCTGCTTGCAGAGGAGGAGGCTGCTCTCTGCTCTCTGCTCAGACCTCCAG 73383
 QY 1215 ProSerThrAlaIleGluGluAlaProGlyLysSerCysGlnProLysAlaLeuCysPro 1234
 Db 73382 CCAGCAGCTGCTATAGAGAGGCCCCCTGGAAAGCTGCCGCCCAAGACCCCTGTGTCT 73323
 QY 1235 LeuThrValSerProSerLeuProArgAlaProValSerSerAlaGlnValPro 1252
 Db 73322 CTAACAGTCAGCCCAAGCCTTCCAGCGGCCCTGTCTCTCTCTGCTCAGGTCCCC 73269

RESULT 12

AC106509/c

LOCUS

DEFINITION

Rattus norvegicus clone CH230-20218, WORKING DRAFT SEQUENCE, 3

unordered pieces.

AC106509

AC106509.4

GI:30580817

HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.

Rattus norvegicus (Norway rat)

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE

1 (bases 1 to 224227)

AUTHORS

Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J.,

Allen, C., Allen, H., Alebrooks, S., Amin, A., Anguiano, D.,

Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,

Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,

Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,

Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,

Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,

Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falle, T., Fan, G.,

Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,

Fraser, C. M., Gabis, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,

Gebregiorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,

Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,

Harvey, Y., Hawlak, P., Hawes, A., Henderson, N., Hernandez, J.,

Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hoques, M.,

Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,

Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,

Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,

Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,

Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,

Lorensuhewa, L., Loulsegged, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okwono, G., Olarpunsaagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wiczyska, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE
JOURNALUnpublished
2 (bases 1 to 224227)

Worley, K.C.

REFERENCE
AUTHORS
TITLE
JOURNAL

Submitted (12-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
TITLE
JOURNAL

3 (bases 1 to 224227)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 13, 2003 this sequence version replaced gi:22857495.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GLBA
Center clone name: CH230-20218
----- Summary Statistics

Assembly program: Atlas 3.0;
Consensus quality: 211258 bases at least Q40
Consensus quality: 214044 bases at least Q30
Consensus quality: 215992 bases at least Q20
Estimated insert size: 220399; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 221791: contig of 221791 bp in length
* 221792 221891: gap of unknown length
* 221892 222909: contig of 1018 bp in length
* 222910 223009: gap of unknown length
* 223010 224227: contig of 1218 bp in length.

FEATURES
source

1. 224227
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-20218"

misc_feature

1. 1752
/note="wgs end extension
clone end:Sp6"

misc_feature

3617..3835
/note="clone boundary
clone end:Sp6
site:BCORI
end sequence: BH334357"

misc_feature

220120..221791
/note="wgs contig"

BASE COUNT 58284 a 52300 c 50849 g 55959 t 6835 others
ORIGIN

Alignment Scores:

Pred. No.: 2.59e-114 Length: 224227
Score: 3064.50 Matches: 984
Percent Similarity: 21.91% Conservative: 24
Best Local Similarity: 21.39% Mismatches: 45
Query Match: 46.40% Indels: 3554
DB: 2 Gaps: 18

US-09-754-997A-2 (1-1252) x AC106509 (1-224227)

Qy 171 ProfileleThrTrpGluLysAspGlnValThrValProGluProArgLeuIleThr 190
Db 111847 CCCCCTCTGACA-----GAAAGAACCATGACGCCCTGACCCCTGCTTTGGTC--- 111800
Qy 191 LeuProlyserTrpLeuLeuGlnIleLeuAspValGlnAspSerAspAlaGlySerTyrArg 210
Db 111799 ATCCGAG----- 111791
Qy 211 CysValAlaThrAsnSerAlaArgGlnArgPheSerGlnGluAlaSerLeuThrValAla 230
Db 111790 -----CAGAGGGTGGCAAGCTGACCTGTCTCTCTATCTTC 111755
Qy 231 LeuArgGlySerLeuGluAlaThrArgGlyGlnAspValValIleValAlaAlaProGlu 250
Db 111754 TCTACAGGGTCTCTGGAGGCTACCGAGGGCCAGGATGTGGTCATAGTGCAGCCCGCAG 111695
Qy 251 AsnThrThrValValSerGlyGlnAsnValValMetGluCysValAlaSerAlaAspPro 270
Db 111694 AACACCACAGTGGTGTCTGGACAGAGTGTAGTTGGAGTGTGTGGCTCCGCTGACCCC 111635
Qy 271 ThrProPheValSerTrpValArgGln----- 279
Db 111634 ACCCTTTTGTCTCTGGTCCGACAGGGTCTGAGTCTGAGGAAAGAGGAGGAGATTG 111575
Qy 279 ----- 279
Db 111574 GGGGAGTGGCAAGGTGTGTACATATCTCACAGGCTCCAGTCAGTCTTACCTAAA 111515
Qy 279 ----- 279
Db 111514 GATGGAAGGAGAGAAATTGACTCCACAAAGTTGTCTCTGCTTCCATGAGCACTCTCAA 114455
Qy 279 ----- 279

Db	111454	CAAGCGCATGTGCACACATTCATGCGATTAATAAGTCTAAAGAGGTTGTGTAAAGCC	111395
Qy	279	-----	279
Db	111394	CAGGGATGAGGGGCACATCTATCCCAAGCACTCAGAGGCGAGAGCGAGATCTCT	111335
Qy	279	-----	279
Db	111334	GAGTTTCAGGGCAGCCGGACTACTCAGAGAAACCTCGCTCGAATAATAGAGGGCAGG	111275
Qy	279	-----	279
Db	111274	GCGGGGAGAGAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACA	111215
Qy	279	-----	279
Db	111214	GACACAGAGAGAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACA	111155
Qy	279	-----	279
Db	111154	GAGACAGAGACAGACAGACAGACAGATAGACAGATGGGAGCAGGGAAGGGAAG	111095
Qy	279	-----	279
Db	111094	GGACAGAGTTGTATTAATATCTGCCGCCACACAGAGGCCATTTAGCCGTCGTTACC	111035
Qy	279	-----	279
Db	111034	CACCTCCCATCTGCTTTCCACCTAGGGATCAGAGTTAGAAGGAACGTGATTCAAGCCAG	110975
Qy	279	-----	279
Db	110974	CGACTGTGTGCAGCTTTGCGGGGTTTCGGGAGTGGCGGTGTAGTAGGATCTTTGTAGTTG	110915
Qy	279	-----	279
Db	110914	GAGAGGGTGTAGGGGACAGTTCCAGGTCCTCCCTTAACCTACTCTCTCTGTGCACCA	110855
Qy	280	AspGlyIysProIleSerThrAspValIleValLeuGlyArgThrAnLeuLeuIleAla	299
Db	110854	GATGGGAAGCCTATTTCACGGATGTCATCGTTCTGGCCGGACCAATCTACTCATCGCC	110795
Qy	300	SerAlaGlnProArgHisSerGlyValTyrValCysArgAlaIleIysProLeuThrArg	319
Db	110794	AGCGCGAGCTCGGCACCTCTGAGTCTATGTCTGCGGGCCCAACAAGCCCGCACCGGA	110735
Qy	320	AspPheAlaThrAlaIleAlaGluLeuArgValLeu-----	331
Db	110734	GACTTCGCCACTGCCGCTGCTGAGCTCCGAGTCTGGTAAAGTCTGGCGCTGTGGGACA	110675
Qy	331	-----	331
Db	110674	TTAAGGACAGGTTTGTGTGTAGAGAGAGAGGGCGATCCTCGAAGACTTAAACTTGAGG	110615
Qy	331	-----	331
Db	110614	GGCACGACATGGGAGGAGTCTGAATTACTGATTGATGGTATTGTTGGAGGACTGGATT	110555
Qy	331	-----	331
Db	110554	CGGTAGCTGCGCGGGCGGGGGTGGGGTGGAGGTGGATGGGCGCAATTGAGAGAGGA	110495
Qy	331	-----	331
Db	110494	ATGGAGGAGGGGTATTTAAGCCCTGGCTTCTGTCTATGACTAGTGAAGCGTGGGAGAG	110435
Qy	331	-----	331
Db	110434	CAGGAGTGGGGTATTAGGGAGGAGATAGTGAATGGTTTCTGAGAGGGTATGAAGC	110375
Qy	331	-----	331
Db	110374	TCAGGGCCAGAAATCTGTAGCTGTAGCAATTTGGAAAAGCTGTAGCAGTAATGTGAGTTCT	110315
Qy	331	-----	331
Db	110314	ATGCCAAACCGACCTACAGCTACAGAGTGAAGAGCCTATCTCAAAACAAACTTATCAAAC	110255
Qy	331	-----	331
Db	110254	AAACAAACAAACAAACAAATCAGCTGGAGTGTGTCTGAGGTGACAGATCAGAAATTCAA	110195
Qy	331	-----	331
Db	110194	GGCCAACTTTGGGCTACAAAGAGATTTCTCAATACAAACAAACAACTGGGGGCCAGGCTTG	110135
Qy	331	-----	331
Db	110134	CTGAAAGCTGGAGTAACCTGCTGATATAGAGACAAATGGGAGGAGGGAGTGAAGTTACA	110075
Qy	331	-----	331
Db	110074	GCCTAGAGCTTTAGACTGGAATCTGCTGGCGCTGGGAGAAAGAGGCGCTGTGTGCCGA	110015
Qy	331	-----	331
Db	110014	ATGGGCGAGGGGTGGACAGAGCTGGTCCGGGTGACTAGGAGAGAGGACCTTAGAGTG	109955
Qy	331	-----	331
Db	109954	AGGCTGGGGCGCGGTGTTGGGACACTCGGGGAGGGTAGGCCCTCTTGTACAGAGTCTCC	109895
Qy	332	-----AlaAlaProAlaIleSerGlnAlaProGluAlaIleSerArgThrArg	347
Db	109894	CTCTCCCTCCAGCTGCCCGACCATCTCGCAGCGCCCGAGCGCTGTGCGGACCCAGG	109835
Qy	348	AlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLeuHisTyr	367
Db	109834	GCCAGCACTCGCGGCTTCGTTTCCCGCGCTCGGGGAGCCAGCACTACGCTGCACATG	109775
Qy	368	LeuHisAspGlyIleProLeuArgProAsnGlyArgValIysValGlnGlyGlyGly	387
Db	109774	CTGCACGACGGGCGCCGTTGGACACCTCGGGGAGGGTAGGCCCTCTTGTACAGAGTCTCC	109715
Qy	388	SerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrTyrGlnCysValAlaGlu	407
Db	109714	AGCTTGTGTCATCAGATCGCTGCGCTGACAGCACTCTGCTTACTACCACTGCTGAGCGAG	109655
Qy	408	AnSerAlaGlyThrAlaCysAlaAlaProLeuAlaValValValGluGlyLeu	427
Db	109654	AACAGCGGGGAACCCGCTGTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG	109595
Qy	428	ProSerAlaProThrArgValThrAlaThrProLeuSerSerSerSerValLeuValAla	447
Db	109594	CCGAGTCCCCGACCTCGGCTCAAGGCAACGCGCTGAGCAGCTCCGCTGTGCTAGTGGCC	109535
Qy	448	TrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTyrGlnLys	467
Db	109534	TGGAGCGGCTGAATTGACAGCGAGCAATTCATTGGCTTCTCTCTTCTACACCAAAAG	109475
Qy	468	AlaArgGlyVal-----	471
Db	109474	GCAAGGGGTAT-GTCTACTCTTGACCA CGGGACCCCAAGTTAGGGCAAGTACGGGATTTGAT	109416
Qy	471	-----	471
Db	109415	TAGTAAAGAAATTCCTTAATGGTGACCTCTGCAGATGAGGGGAAGGATGAGAGCAATCC	109356
Qy	471	-----	471
Db	109355	TTCAGGGAAGTGTCTTTCAGTCTAGGATTAGATTATAGAGTGTACACACACACTCTCTT	109296
Qy	471	-----	471
Db	109295	CCCGGAGAGAAATGCTTTTACAGGGTTAAAGGAGCTGTGCTCCACCCCGAGCTAACTGCAGGA	109236

QY	471	-----	471	Db	108155	AGGGGTCCTCCCTAGTGGGGAGTTGTCTAGTGGGAATGGGCTTAATGGGGTATG	108096
Db	109235	GGGACACTTAGACAGCGAGGAACTTCCTGGCTATCTCGAATTAATACATGGATA	109176	QY	569	-----	569
QY	471	-----	471	Db	108095	GTAAGCTCTGGAAGAGTGCCAGAGACAGATAGATAGGTGGTCAAGCCGGAAG	108036
Db	109175	GACTCGATGGCATTTATATCTCCACCTTAGATTGTCTGTGACATGTAGCTATACATAG	109116	QY	569	-----	569
QY	471	-----	471	Db	108035	ACTGAGTTCAATCTCAGACTTCATGTTTAAAGTGGGGGGAAGGCTGTCAAGTGGA	107976
Db	109115	ACCAGTAGAATCCCCAGTCTCCAGCTGTGCATAGGGAGTGGGCTGTGCAGTCCCTTT	109056	QY	569	-----	569
QY	472	-----	482	Db	107975	GTGCTGTTTGTATCCAGTGCAGGGAGGTGGAGACAGGAGACACCTTGAGGTCATT	107916
Db	109055	GTCCCCCGTGACACCTCAGAGTAGACAATGTGGAGTACAGTTTGGAGTAACAAT	108996	QY	569	-----	569
QY	483	AspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPheTyrVal	502	Db	107915	GGGCTAAAGCCAGAGTTCCAAACCAAGTGCAGAGACCTGTGCAAGAGAAGGTGACAAT	107856
Db	108995	GACACACAGAGCTGCAGTTCCGGACCTGGAAACCAACACAGGATTACGAATTCACGTG	108936	QY	569	-----	569
QY	503	ValAlaTyrSerGlnLeuGlyAlaSerArgThrSerProAlaLeuValHisThrLeu	522	Db	107855	ACCAAAGAAATGACAGAGGTTACATCTGACCTTACATGTGTGGCACACACATGTAC	107796
Db	108935	GTGGCTACTCCAGCTAGGGCCAGCGGACCTCCAGCCAGCCCTGGTGATACACTG	108876	QY	569	-----	569
QY	523	AspAsp-----	524	Db	107795	CTGCACATATGTGAAACTCACACTCATGTGACACACACAGAAACAGTGTGTTGGGGCA	107736
Db	108875	GACGATGGTAGGACCTCTGAACTGTCAGTGCCTGGGACCCAGAGACACTCCCAAGCAGG	108816	QY	569	-----	569
QY	524	-----	524	Db	107735	GTTGAAGAAAGGGAAGACAAAGAGGGTCACTAGCTTGGGCA TGGTAAGAAGTATAATC	107676
Db	108815	GACAGTGTGGCCCAAGGCCCTCTCCCTGCCCTTGTAAATGTCGTGTCGCCCTATAAGCA	108756	QY	569	-----	569
QY	524	-----	524	Db	107675	ACTGGAGCCTTGGACTCAGTGTCCCATCTTTTGAACCTCTTAGGTTTTTCTTTGTTT	107616
Db	108755	ATGGGCAGAGTTCTTCTCCAGTGTAAATATGGGAAACTGAGGCTCAGAAATGAAC	108696	QY	569	-----	569
QY	524	-----	524	Db	107615	GTCCTTTAAGATTATTTATTTATGTATATAGTGTCTTTATTTGCATGACACCTGCAG	107556
Db	108695	GACTTGCCAAAGCTCAGCAGACAGCTGGGGCTTGAGCCAGCTCTCCAGTCCCTCTACAG	108636	QY	569	-----	569
QY	524	-----	524	Db	107555	GCCAGAGAGGGCATCAGATCCCATTTACAGATGTTGTGAGCCATGTTGTTCTGGG	107496
Db	108635	TGCAGCCCTTCTTCTTAATGCTCATGTACCATCCAGCCATTGACTCGAAAGCTTCTGGGC	108576	QY	569	-----	569
QY	524	-----	524	Db	107495	GATTGAGCTCAGGACTTTTGAAGAGCAGCCAGTGCCTTAACTTAAGCCATCTCTCC	107436
Db	108575	CACACTGTTCACTAGCTTGGGATAGATGTCTCTCATCTCTAGCCTAGGGCTC	108516	QY	569	-----	569
QY	524	-----	524	Db	107435	TGCCCCAGAAACCTTCTAGATTAAACACAGAGAGGGGGAGGCTGCAGGAGTGAC	107376
Db	108515	CCCCGAAAGGCTCCCTTCTCCCTGTGATCCCGGCCCCCGGCTACTTCTGACTTGTG	108456	QY	569	-----	569
QY	524	-----	524	Db	107375	ATGGGCATGGCTGGTCTCTACTGGGGAGATGATGCTGTTCTGGGAAGCGATGTGTT	107316
Db	108455	TGTGACCTGGACAGTCTCTGACCCTCTGACCTTGTGTTTCCCTGTGGTGAGAGGAGT	108396	QY	569	-----	569
QY	525	-----	536	Db	107315	TGAAAGGAATGGAGAGTATGGAAGAAAGAGTGAAGAGGTGAGGTTAAAGATAAGAG	107256
Db	108395	TCATGCCAATGTTCTTTGTCTCAGTCCCGCAGCGACCCAGCTCTCCCTGTCCAGC	108336	QY	569	-----	569
QY	537	ProAsnProSerAspIleArgValAlaTrpLeuProLeuProSerSerLeuSerLeuGly	556	Db	107255	AGAAGCAGGAGGTGACCCCAAGGCGAGTCTGAGGCTCTCTGACACCAGGTACTGTTAGT	107196
Db	108335	CCCAACCTTCGACATCAGGGTGGCATGGCTGCCCCCTGCCCTCCAGCCTGAGCAATGA	108276	QY	569	-----	569
QY	557	GlnValLeuLysTyrIleGluTyrGlyLeuGlyLys-----	569	Db	107195	GCTCTTAGCTAATCATCAAACTGTGTACAAAGGGTGTGAGCTGGCTGCTATTGGTTCA	107136
Db	108275	CAGTGGTGAAGTACAAAGATAGGTACGGTTCAGGGAAGAGGTGAGTGGGCTCAGGC	108216	QY	569	-----	569
QY	569	-----	569	Db	107135	TAGGGAGGAGACCATAGCCGTGTCTTCCATGTTGAACCAAGCCAAATGGTTTCAAC	107076
Db	108215	AGTCAGAGATGACTGTGGAAATGGGGTCTCCCTCCTAGGTAGGGGAGTGTCTCAGTGGG	108156	QY	569	-----	569
QY	569	-----	569				

QY	740	-----	740
Db	104857	CCAGCCAGCAGCTGTGTGATGCCACAGCCACCCAGCACACCTCTCTGGCGAGGAGTTT	104798
QY	740	-----	740
Db	104797	TCCTGTCTAGACATGCCCTCATTTCTCAAGCTGAGCAGAGCAAAATAGGGAAGTTTGA	104738
QY	740	-----	740
Db	104737	AAGAGAACTCCATGTCGTGATGCGTGATTCCTGGCTTAATTTTAAATCTTAGATAGTTT	104678
QY	740	-----	740
Db	104677	TTAAGATCATCCACTTATACCACAGGATTCATCAGATTTGTCCCATAGTTTCTTTTAA	104618
QY	740	-----	740
Db	104617	GAITTAATTTATTTATTTATATATACGAGAGTACACTGTACTTGTCTTTTGGACACCC	104558
QY	740	-----	740
Db	104557	AGAAAAGGCATAGGATCCCACTACAGATGTTGTGAGCCACCATGTGTTTGTGGGATT	104498
QY	740	-----	740
Db	104497	TGAACCTCAGCTCTCAACGACTGAGTCATCTCTCCAGCCCCCATCCCATAAATCTTAACTAC	104438
QY	740	-----	740
Db	104437	AACACAGTGGTTTGTCTACCTTGACAATCCATCTGACAGTCAGAAGCCAAACCCCTGGGT	104378
QY	740	-----	740
Db	104377	TCTAAATTCAGCTAGGTGATATCTCACTCAGCAAGTGTCTTACTAAGTATCTACCACGT	104318
QY	740	-----	740
Db	104317	GCCAGCTCTGTTCTGGAAATATAGCAATGAGCAGCATGAATAGATATCTGTTCTCTGGG	104258
QY	740	-----	740
Db	104257	AAITTTGTTTCATGCTTATCTAATAACAATTTTTCATCACCANAATAGTATCCACAGAGTT	104198
QY	740	-----	740
Db	104197	TTCTACCCACGCTGTAGTATTAACTAATCCAACTTCTAAGAACCTCAGAGGGGACC	104138
QY	740	-----	740
Db	104137	AGTGAGATGGTTACAGCAAGTGAAGGTGCTTGGTACAAGCCAGATGACCTGGGCTTGTTC	104078
QY	740	-----	740
Db	104077	CCGGCACTCCCAAAATTTGCTCTCTGACCCCAATGATATAGTGAGCTCATTTCCCTTC	104018
QY	740	-----	740
Db	104017	CCTCCACACAATAATCATGTGTGATCGAGAAGCAGCAGTGATTTGAGAATTTTAAAGAA	103958
QY	740	-----	740
Db	103957	ACAAATGTAAAAAAGGAGAAAGAGAGGAGAAAGAAAGAAAGAGCAGCTCGAGAT	103898
QY	740	-----	740
Db	103897	GTTGGTGTGTCTGTCAGATCCATATATAAGATPAGAGACAGAGCTACAGAGAAATCCCA	103838
QY	740	-----	740
Db	103837	ACTTCCAAAGCCACATAGTCAGTAAGTGGCAGAGAGGAGTTTCAACCAAGGAGCCT	103778
QY	740	-----	740

Db	103777	CGTTCCAGAGTCATGSCATGCTTTCTCTCTCTCATATTTTCATCAGTCATGGATGTGTCTAC	103718
QY	740	-----	740
Db	103717	AGTCACTCTATGTGTGACCTTAGCAAAAATTCATTTCAACTCTGGGAGCCTCAGTTGCA	103658
QY	740	-----	740
Db	103657	GCAGGATTAACATCACCCCACTTACAGGATCAAAAGAGAGGCCCGGTGCCAGCTATGGTA	103598
QY	740	-----	740
Db	103597	TCCACAGCTGTAAACCCCTGTGTCTAGAGGATCAGAGCAAGAGCTAGCCTGGGCTACTAC	103538
QY	741	-----	745
Db	103537	ATAATGAACATCTGCCCATCTCTCTTACACTTGGATTTTCCCTCTAGATCTGCCTATCC	103478
QY	745	-----	765
Db	103477	lnArgGlyProLeuProAlaHisAlaHisAlaGluSerAsnSerSerThrSerI	103418
QY	765	-----	785
Db	103417	leTrpLeuArgTrpIysLysProAspPheThrThrValIysIleValAlaAsnIyrThrValA	103358
QY	801	-----	801
Db	103297	AAGGGAGGGCAACTCGAATAGTAATACTAGCGGGCTCAGCTATGGCTAGGGCACACCCAG	103238
QY	801	-----	801
Db	103237	AATCTACACAGATGTCTCTTTAGAACTGTCTGACAGGGAGTAACAATAGGGAGACTTGT	103178
QY	801	-----	801
Db	103177	CTGGTCCCAACCCCTAATATGACCGAGTGTGTGTCAACCTTTGGTGTGAGTGGGATGCCATC	103118
QY	801	-----	801
Db	103117	CATTGGCTGGACCCCTGCCCTCCCATCATGGAGAGATAGCTTTAAGGAACCTGCTGTCTG	103058
QY	801	-----	801
Db	103057	ATTGAGAAGAGCCCTGTGAAAGCCCAACAGAGAGATAAACTAAGCCAAGGATAGGTAGAT	102998
QY	801	-----	801
Db	102997	ACCTCAGAAGTGACTGGTCTCCCCACCCCGGTAAAGAAATACTTAGATTAATTCAGGAA	102938
QY	801	-----	801
Db	102937	GGCATCCCCGAGAGAGAGAGGTCTGCTTTGGAGATTAGACATGGGAATACCATATGTCCG	102878
QY	801	-----	801
Db	102877	ACCTCTCTGTCTGTTTCCCTGGCTGTGTGTGGGAGGGAGGTAGTCTCCCTGAGTAGCAG	102818
QY	801	-----	801
Db	102817	CACCTTCAGGTTGGGGCCTGGCTGGAGCGAGGGCTTATGTCTGGGGTTCTGCCTGGCT	102758
QY	802	-----	807
Db	102757	TGAGGCTGTCTGAGGGCACTAGTCTCTCTTTCTCCGGAGAGCTCTGGAGAGACATC	102698
QY	808	-----	827

QY	965	yLeuLeuCysLeuAlaCysMetCysAlaGlyLeuArgInserSer-----	981	Db	99400	CCACTTTCTCTGCAATTAAGGAAGAGTCCACCAGGTTTGGAGCTTTCGGATTTTGGGATT	99341
Db	100479	CCTTCTCTGCTCTGGCTGCAATGTGTCTGCTGCGACGAGCTTCGCACAGGTGAAC	100420	QY	1060	-----	1060
QY	981	-----	981	Db	99340	TTGCACATCAAGTGTATTAGGCTTATGATCCGCCCTTCTCTATCTCTCCACCTCAGTT	99281
Db	100419	GAGGAAACCATTTAGAGCTAAGAGATTACCAGGAAGGGCTGCTCTGAAGCCCAATATA	100360	QY	1060	-----	1060
QY	981	-----	981	Db	99280	CTTACCCCTGCTGGAGGGGGCTTTAGCGGCAACCTATCTGGAGATAACAGGGGGCCC	99221
Db	100359	GTTAGAGCTATGAAATGAGTGGAGTTTCACTGCTACTTTTGGCCCTCATGTGAGGCCCTG	100300	QY	1061	-----	1061
QY	981	-----	981	Db	99220	ACAGGTCTGGACCCCTGCATACCATATCTTTCTTCTCTGCGAGATCTCTCTGGCTC	99161
Db	100299	GGCTCCATCATCAGTAGCACAAAAACAGTCGATAATTTATTATTAGCATTTTGGATG	100240	QY	1064	lnAlaGlyClyProAsnTrpAlaGlySerTrpAlaGlyCysGluLeuProGlnGlySerG	1084
QY	981	-----	981	Db	99160	AGGCAAGGGGACCAAAATTTGGGAGGCTCTCTGGGAGGCTGTGAGCTGCCACAGGGAGTG	99101
Db	100239	AAGGATTTGGGGGCTTGATCACATGAGACTGTAGATAAACCAGGGGAGCTTACCCCCC	100180	QY	1084	lyProArgProAlaLeuThrArgAlaLeuLeuProProAlaGlyThrGlyGlnThrLeuL	1104
QY	981	-----	981	Db	99100	GTCCAGGCGGCTTAAACCGGGCTCTGCTGCTCCAGCCGMACTGGGCGAGACACTGC	99041
Db	100179	CCTACCCGTCCTCCCTCACCTCACAGCAGGTGAGTACTGTGGGCTTTTCTCATAGCTT	100120	QY	1104	euLeuGlnAlaLeuValTyrAspGlyIleLys-----	1114
QY	981	-----	981	Db	99040	TGCTGCAGGCTCTGGTGTATGATGCCAT-AAAGGTTGGTCACTGTGTGAGAGAGAGG	98982
Db	100119	CCCTGCTGGAGCTTAGATGGAAGCTCCACGTCGCCATTCAAGGAGGGTGCCATTG	100060	QY	1114	-----	1114
QY	981	-----	981	Db	98981	GAGGCAAGGGGAAATCCCTATGCAATTTCCAGGGGACAGAAAGTATTTGGAGACCTTG	98922
Db	100059	AGGGATGGACTGTATTCTCTCTTCCGTTGATCACACACACAGGAAGCCAAATCC	100000	QY	1114	-----	1114
QY	982	-----	HisArgGluAlaLeuProGlyL	Db	98921	TAGCTTCCAAAGGGAGTGGTGACAGGAGCCAGTTTGGAGCAGTCTCTTCTTTACCTG	98862
Db	99999	TCACATCTTGTCTGACCCATAGCTTTGTTCTGCTGCTCACAGGAAGCCCTCCCGGAC	99940	QY	1114	-----	1114
QY	989	euSerSerGlyThrProGlyAsnProAlaLeuTyrThrArgAlaArgLeuGlyProP	1009	Db	98861	GACCTAAATGGCTCATAGTAGACTTACAGAGTATCTGAAGGTGGCCCTGACCCCTCAGATA	98802
Db	99939	TGCTCTCCAGGGCACCGCTGGGNAACCCAGCATCTTACTCAAGAGCTCGGCTCGGCCCC	99880	QY	1114	-----	1114
QY	1009	roSerValProAlaAlaHisGluLeuSerLeuValHisProArgProGlnAspTrpS	1029	Db	98801	CGGTCAAGCCCTGCTCTCTTTATGAGAGTCTGTGCTTAGTACTTTCTAGTAGCCTGAG	98742
Db	99879	CCAGTGTCCCGCTGCGCATAGTTGGAGTCCCTTGTGATCCTCTGCTCCCGAGACTGGT	99820	QY	1114	-----	1114
QY	1029	erProProSerAspValGluAspIysAlaGluValHisSerLeuMetGlyGlySerV	1049	Db	98741	TGAGGGCAACCGGGCATCTGTACTTCTTGCTCTGCAAGAGCCTGAGCANTCTCAGGCG	98682
Db	99819	CCCCACCGCTCGGATGTGGAAGACAAGCTAAAGTGCACACCTTATGGGTGGCAGCG	99760	QY	1114	-----	1114
QY	1049	alSerAspCysArgGlyHisSerLysArgLysIle-----	1060	Db	98681	GAACTGTGCTTGTGTGTTGTCCTCAATGGGGAGTTCAGCCCTCTCCACCCCTACATGTGTTGAT	98622
Db	99759	TTTCAGATTGCCAGGCCACTCCAGAGAAAGGT-GAGAGGTGAGACTGGAGTCTCGGTA	99701	QY	1115	-----	1115
QY	1060	-----	1060	Db	98621	CATTTCTGAAAGACCATCTCAGAGACGCTCACCTTCTGTTTCTCCACCTCTCTGCCAGACA	98562
Db	99700	GGTCAGAGGCTCACAGGACCCATTTCCAAATCAAACACTAGGGGGCGCCAGAGCAAC	99641	QY	1116	snGlyArgLysLysProSerProAlaCysArgAsnGlnValGluAlaGluValIleValH	1136
QY	1060	-----	1060	Db	98561	ACGGGAAAGAGACCGCTCCACGCTCGAGGATCAGTGTGGGGCGAGAGTCAITGTCC	98502
Db	99640	TTCTGAGCGGACGACACACAGGCCCTGAGCTATCTAGAGAAAGTGTCTGTGCCCCAC	99581	QY	1136	isetaAspPheGlyAlaSerLysGlyCysProAspLeuHisLeuGlnAspLeuGluProG	1156
QY	1060	-----	1060	Db	98501	ACTCCGATTTAGTGCATCCAAAGGATGCTCTGACCTCCACCTCCCAAGACCTGGAGCCCG	98442
Db	99580	CTGGAGGGCACCCAGAGCTCGGAGCACCCATCTTCTCTCACTAACCTAATATGGAGAA	99521	QY	1156	luGluProLeuThrAlaGluThrLeuProSerThrSerGlyAlaValAspLeuSerGlnG	1176
QY	1060	-----	1060	Db	98441	AGGAGGCCCTGCTGCAGAGGTTCTGCCTCCCACTGCTGGGTCTGTGGATTGTCTCCAAG	98382
Db	99520	AAGAACTGTGCTTGTGCTGGCCAAATAGGCGTGGCCCGCCAGGATGAGGTGAGATACC	99461	QY	1176	lyAlaAspTrpLeuGlyArgGluLeuGlyCysGlnProThrThrSerGlyProGluA	1196
QY	1060	-----	1060	Db	98381	GAGCAGCTGGCTGGCGAGGAGCTGGAGAGTGCACACAGCAACCTCTGGGCCACAGA	98322
Db	99460	GTAGGCCAGGTTCCCTGGTGGCACACCTTGACCAAAATGGACAGTCTCTCCCATCTCT	99401	QY	1196	rgLeuThrCysLeuProGluAlaAlaSerAlaSerCysSerCysSerAspLeuGlnProS	1216
QY	1060	-----	1060				


```

misc_feature      /note="assembly_fragment"
2886..4859
misc_feature      /note="assembly_fragment"
4960..6416
misc_feature      /note="assembly_fragment"
6517..8739
misc_feature      /note="assembly_fragment"
8840..10868
misc_feature      /note="assembly_fragment"
10969..13737
misc_feature      /note="assembly_fragment"
13838..15329
misc_feature      /note="assembly_fragment"
15430..17938
misc_feature      /note="assembly_fragment"
18039..19981
misc_feature      /note="assembly_fragment"
20082..23587
misc_feature      /note="assembly_fragment"
23688..27447
misc_feature      /note="assembly_fragment"
27548..30161
misc_feature      /note="assembly_fragment"
clone_end:T7
vector_side:left
30262..34737
misc_feature      /note="assembly_fragment"
34838..40101
misc_feature      /note="assembly_fragment"
40202..44039
misc_feature      /note="assembly_fragment"
44140..49950
misc_feature      /note="assembly_fragment"
50051..55555
misc_feature      /note="assembly_fragment"
55656..60496
misc_feature      /note="assembly_fragment"
clone_end:SP6
vector_side:right
60597..68544
misc_feature      /note="assembly_fragment"
68645..76940
misc_feature      /note="assembly_fragment"
77041..85480
misc_feature      /note="assembly_fragment"
85581..93905
misc_feature      /note="assembly_fragment"
94006..108063
misc_feature      /note="assembly_fragment"
108164..123496
misc_feature      /note="assembly_fragment"
123597..139965
misc_feature      /note="assembly_fragment"
140066..176282
misc_feature      /note="assembly_fragment"
2608 others
BASE COUNT 44272 a 43276 c 42765 g 43361 t
ORIGIN

```

```

Alignment Scores:
Pred. No.:      9.29e-94      Length:      176282
Score:          2557.00      Matches:      877
Percent Similarity: 20.85%      Conservative: 45
Best Local Similarity: 19.83%      Mismatches: 111
Query Match:      38.72%      Indels:      3397
DB:              2          Gaps:      18

```

US-09-754-997A-2 (1-1252) x AC068507 (1-176282)

```

Qy      195 LeuLeuGlnIleLeuAspValGln-----AspSerAspAlaGlySer 208
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      167329 CTCATGGAACCCCTGCCAGTGAAGGCTTAATCCCTTACAGATGGAGACAGCGGGCT 167270
Qy      209 TyrArgCysValAlaThrAenSerAlaArgGlnArgPheSerGlnGluAlaSerLeuThr 228
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db      167269 CAGAGAGTCCCGACGAGAGAGAAATGAGATTCGACCCAGGAGCGTGTGTTGCTTT 167210
Qy      229 ValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAspValValIleValAlaAla 248
Db      167209 ATCCGGGTGCAGAGTGTCTCGAGTTCGGAGACATGAGGGAAGGCGAGTCCCTAGCGTC 167150
Qy      249 ProGluAenThrThrValValSerGlyGlnAenValMetGluCysValAlaSerAla 268
Db      167149 CCTCCAGC-----GGCTCTCCC 167132
Qy      269 AspProThrPropheValSerTrpValArgGlnAspGlyAspProIleSerThrAspVal 288
Db      167131 TCTCCTTCCCCA-----GACGGGAAGCCCATCTCCACAGATGTC 167093
Qy      289 IleValLeuGlyArgThrAenLeuLeuIleAlaSerAlaGlnProArgHisSerGlyVal 308
Db      167092 ATCGTCTCGGGCGGCACCACTACTAATTCACAGCCGACGCTTCGGCAGCTCCGCGCTC 167033
Qy      309 TyrValCysArgAlaAenLysProLeuThrArgAspPheAlaThrAlaAlaGluLeu 328
Db      167032 TATGTCGCGCGCCACCAAGCCCGCACGCGGACTTCGCCACTGCACGCGCTGAGCTC 166973
Qy      329 ArgVal----- 330
Db      166972 CGTGTCTGGGTAGGCGGGCCGAAGGGCGGAGGCGGGTGTAGCTGGGGCTGGGGTGA 166913
Qy      330 ----- 330
Db      166912 ATAGACAGGGGAGGTGACCGGGTGGGTGTCTGGGGGACATCAGCGCGCTCTCTGGAG 166853
Qy      330 ----- 330
Db      166852 GACGAATCAGGGAGGAGTCCCTAGTAAGGCGGAGGTGGGACGAGGATGGGACTTGGT 166793
Qy      330 ----- 330
Db      166792 AACTGGAGGACGAGTGCAGGGCGGTGAGTCTGGGATGGAGCGAGGGGTGATTAAGGC 166733
Qy      330 ----- 330
Db      166732 CTGGACGGGTGATTAAGCGGGAGTGTGATCGGGGTATTGTGGGGATAGTGGAGTGA 166673
Qy      330 ----- 330
Db      166672 CTGAGTGGTACTAAGAAACTGGGGGCGAGGTAGATCTGTTAGTGTGGACATATGG 166613
Qy      330 ----- 330
Db      166612 AGGGCAGGAGGAAGGAGATGGGTAATTTAGGGGTGGGTAGATCCGGGAGGAGCAAT 166553
Qy      330 ----- 330
Db      166552 ATTTAGGGAGGCTGAGGAGGCTGGGGACAGTTTGGGAGAGAAAGAGAATGTCGTTG 166493
Qy      330 ----- 330
Db      166492 TGGGGCTGGGGTAAATGGGGAAGAAAGGGTAGGTCTGCGCTGAGCCGAGTAGTTGG 166433
Qy      330 ----- 330
Db      166432 AAAGGGTGAATCAGGAGAGGTTAAAGTGAATTCGAGGGGAGCGAGGGCGGTGGGG 166373
Qy      330 ----- 330
Db      166372 GGCAGGCTTGGGGTACTGGGGGGGGGCGAGAGACCTTAATTTGGGGATCCCGGCTCAGC 166313
Qy      331 -----LeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerAr 345
Db      166312 GTGCGCCCTCGCGCCCTAGCGCTCCCGCATCATCTCAGCGCCCGGCGCTGTGCGG 166253
Qy      345 gThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLe 365
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```



```
Db 166252 GACGGGGGAGACACAGCGCGCTTCGTGTGCGCGCGCTCGGGGGAGCGCGCGCCAGCGCT 166193
Qy 365 uHieTriLeuHisAspGlyIleProLeuArgProAenGlyArgValIysValGlnGlyI 385
Db 166192 CGCTGTGGCTGCAACAGCGGGCGCGCTCGCGGCCAACAGCGCGCTCAAGGTCACAGGGCGG 166133
Qy 385 yGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrTyrGlnCysVa 405
Db 166132 CGGTGGCAGCTGTGTCATCACAGATCGGCTCGCAGACCGCGCTACTACCAAGTCGGT 166073
Qy 405 lAlaGluAenSerAlaGlyThrAlaCysAlaAlaProLeuAlaValValValArgG 425
Db 166072 GGCTGAGAACAGCGGGGAAATGGGTGCGCTGCGCGCTGCTGGCGGTGGTGGCGGA 166013
Qy 425 uGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerValle 445
Db 166012 GGGGCTGCCAGCGCCCCACGCGGGTCACTGCTACGCCACTGAGCAGCTCGCTGTGTT 165953
Qy 445 uValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTy 465
Db 165952 GGTGGCCTGGGAGCGGCCGAGATGCACAGCAGCAGATCATCGGCTTCTCTCCACTA 165893
Qy 465 rGlnLysAlaArgGlyVal 471
Db 165892 CCAGNAGGCACGGGGTAC-GTCGGCTGGGGGGGATGSCACAGCGCAGGCCCGGGCGGG 165834
Qy 471 471
Db 165833 GCGCAGCCAGAAGGGTTTCAGTCCGATGATCGAAGGGTTGCTGTATGGGTACCTCTGCG 165774
Qy 471 471
Db 165773 GCAGGATGGACTGGGGTAGGAGCCCATAGGAGAGACCCGCATAGTATGGTAGGTATGG 165714
Qy 471 471
Db 165713 ATTAGGCAACTGGCGCTCGAAGGGGCCACTGTCTTCCTCAGTCATGTTTGTGAGAGAGGA 165654
Qy 471 471
Db 165653 GGGAGAGGACGTTCTCTGGAGCCACCTCCATCCACAGTTAACTAATCTGGAGAGGGGGTCA 165594
Qy 471 471
Db 165593 TTTCAGAAAGACATCCCTTAACAGTGGAGGAATTAATACACTGAGACAGAGAGCTGAGG 165534
Qy 471 471
Db 165533 GAGTTTATACACTCCAGCTCAGATTATCTGGGATGATTTAGCTATAAATAGACACAGGA 165474
Qy 471 471
Db 165473 CAATATGAAGGTTCTCTATACAGACTCCAGCTAGCCAGCAGCGGGGGTGGCACTATGAC 165414
Qy 472 472
Db 165413 AGTCCCTTGTCCCGGTGTACACGAGCATGGCAATGTGGAATACCAAGTTTGCAGTGAA 165354
Qy 481 nAsnAspThrThrGluLeuGlnValArgAspLeuGluProAenThrAspTyrGluPheTy 501
Db 165353 CAACGACACCCAGAACTACAGGTTTCGGGACCTGGNACCCCAACACAGATTTATGAGTTCTA 165294
Qy 501 rValValAlaTyrSerGlnLeuGlyAlaSerArgThrSerProAlaLeuValHisTh 521
Db 165293 COTGTGGCCCTACTCCCACTGGGAGCCAGCGCACCTCCACCCCGCAGCACTGGTGACAC 165234
Qy 521 rLeuAspAspVal 525
Db 165233 ACTGGATGATGG-TAGGGCTCTGAACCTCGCAGTGGCGAGCTTGGGGTAAAGAGAAATTC 165175
Qy 525 525
Db 165174 CTGGGAGGGGAGCTCACCCCTCTCTTTACTAATGTAAGGCATTTTATTACCTCATATC 165115
```

QY	570	-----	570
Db	164034	GGTGGCCCTCAGGGCTCTGGAAGCCAGGCGTTCTTGGCTGGCCCTTCATGTGACGCTCTTG	163975
QY	570	-----	570
Db	163974	TGCTGCAATCCCGTGTGTGTGTGGCAGGGATGTGTCATTGACAGCTCTCGGGGAA	163915
QY	570	-----	570
Db	163914	GTGGGGAAAGCCTGCCATCATTTCTCCCTGCAGCTGAGCCGAGGCTGATGGCCTGCATCAG	163855
QY	570	-----	570
Db	163854	CCAGAGTTGGAGGCTCAAGCGTCAACCCCTGGAGCCATAGAACTGCCAGGCACTGAGGG	163795
QY	570	-----	570
Db	163794	GTCAGTGCTGAGAAACGCTCAGGCTCCACGGGCACGGAGAACCGGATGTGAGGCGGA	163735
QY	570	-----	570
Db	163734	AGAGCCATAGGGGGTTTAGGATCCTGTTCTCAGTGGGTACCCAGCAGCCACCTT	163675
QY	571	-----	586
Db	163674	TTCTGTCCCTGCAGATCAGATTTTCTTACTGAGGTGCGAGGAAATGAGACACAGCTTAT	163615
QY	586	rLeuAenSerLeuGlnProAenLysValTyrArgValArgIleSerAlaGlyThrGlyAl	606
Db	163614	GCTGAACCTCGCTTCAGCCAAACAAGGTGTATCGAGTACGGATTTTCGGCTGGTACAGCAG	163555
QY	606	aglyTyrGlyValProSerGlnTrpMetGlnHisArgThrProGlyValHisAenGlnSe	626
Db	163554	CGGCTTCGGGGCCCCCTCCAGTGGATGCATCACAGGAGCCCGATGTCACACACAGAG	163495
QY	626	rHis-----	627
Db	163494	CNATGTACGGGTTAGAGGAGACAGACAGTGTGTGTGACTGTGTGGTGTGCT	163435
QY	627	-----	627
Db	163434	GTGTCTTGGACCCAGCGCCCCCCTGGCCCGCCCTCCATCAGAGACCTTCCATG	163375
QY	627	-----	627
Db	163374	GCTCCCTCGCTTGATGGGTGAGTGTAGGCTTAGGTGCAGAGGTGAAGGCTGGGATCTGG	163315
QY	627	-----	627
Db	163314	TGTCCCTTCGATTTGCTGTGTGACCTTGGACAAACCATTTCTGAGGCTTCATTTACTGCC	163255
QY	627	-----	627
Db	163254	AGGAATGGCATGAGGTTGAATAATGTTTGAACCTAAAGTGTAGTACATAAAACTCA	163195
QY	627	-----	627
Db	163194	TTGAATTGTAGTCTTCGAGGATGATGGTGTATGATGATTTTGTGAGAGGCCCAA	163135
QY	627	-----	627
Db	163134	TAGAGAGTGATTAATGCTGACCAACCGGTCTCATGCTCTCTATGGACACTAGGTA	163075
QY	627	-----	627
Db	163074	AATTACTTAATCTCTGAGCTTCAGTTTCTTATCTGTAAGGAAGTTGTAGTATTTT	163015
QY	627	-----	627
Db	163014	CACATTTTAGAGGTTGCAAGGATAAAGTTATAATAACTACACAAAGTGCATAGGACAG	162955
QY	627	-----	627

Db	162954	TGTTTGACACATTTGTAAAGCTCTCTCACTCTCAAGTGGGTCTCCCTACTCTAAATGGTGGCAT	162895
QY	627	-----	627
Db	162894	CCCTCTTCTTCTGGCCTCAGCAACCAACCCCACTGCCCCAGTGTGTGATCCTTCA	162835
QY	627	-----	627
Db	162834	AGACTTCACAGAGCCGATTTCTGTGCCAGGTGCTATCAGCCCTACTCCAGAGGTGCTCCTT	162775
QY	627	-----	627
Db	162774	GCTGACCCCTTTTCAAAATGCCAGCGGACACCCCTGCATTTCCCTGCATCTCTGTGCTGTGA	162715
QY	627	-----	627
Db	162714	GAACAGTGAAGTTACCTCTAGACCAAGCCAGGACTGGCACTTAGAGTGTGTGGCTGAGTG	162655
QY	627	-----	627
Db	162654	GTGATGTGTTGGAGATGCTGGCAAAAGGCTCTCACTGAATGCCATCTTCTGTCTCTT	162595
QY	627	-----	627
Db	162594	CCTCATGCTCTTCCCTGTCTCTCTATCCTGGCCCTCTTCTCTCATTTCTGCATTTCTGTT	162535
QY	628	-----	636
Db	162534	CCTTAATTTTCCCATCTCTCTCTGCCCCCAGTCCCTTTTGGCCCTTCAGAGTTGAA	162475
QY	636	sValArgAlaLysMetGluSerLeuValValSerTrpGlnProProHisProThrGln	656
Db	162474	GGTGCGGAAGAATGAGTCTCTGTGTGTATGTCAGCACCCTTCACCCCAACCA	162415
QY	656	nileSerGlyTyrLysLeuTyrTrpGlyGluValGlyThrGluGluGluAlaAspGlyAs	676
Db	162414	GATCTCTGGCTACAACTATATTGGCGGAGGTGGGGCTGAGGAGGCCCAATGGCGA	162355
QY	676	phArgProGlyGlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLysLys	696
Db	162354	TGCGCTGCGAGGGGCGGTGGAGACAGGCTTTGGGATGTGGGGCTGTCTCCGCTCAAGAA	162295
QY	696	slsValLysGlnTyrGluLeuThrGlnLeu-----	706
Db	162294	GAAGTGAAGCAGTATGAGCTGACCCAGCT-AGGTGAGAGGGGCTGGGGGGAGGGGAGG	162236
QY	706	-----	706
Db	162235	GAGGATGTAGCTTCAGTGGGACTGGGCAAGTTGTGTGAAGAGTTAGAGATGGATGC	162176
QY	706	-----	706
Db	162175	TCTCTGTGACGGGAGAGTGACCCAGATGTCCCTTAGCCATTAAAGCCTGTAACTTC	162116
QY	707	-----	725
Db	162115	CCAGTCCCTTGGCGCTGTACGAGTGAAGCTGTGGCTTTTCAACAACAATAGGATGG	162056
QY	725	YTrAlaAlaValTrpLysGlyThrGluLysAlaProThrProAsp-----	741
Db	162055	CTATGACGAGTGTGGAGGGCAACCGGGAAGGGCCCGGCACCAGG-TGAGGGGGGTG	161997
QY	741	-----	741
Db	161996	CGGGAAGAAGGGGGAGGCTCAGGTGAGGCTCTGCTAGCGCTTCTCTTCTTTTCC	161937
QY	741	-----	741
Db	161936	TTTGGGCCAGCTGTGTGAGCAGGTACTTTTAGGCTCAGCTCAGTCAATGCCCAAAACCATCT	161877
QY	741	-----	741

Db	161876	CCTAGGATCAGGGGCTCTCTTTTAGCTCCAGAGAGCCTCTGTCAATGAGAGAGCTTGT	161817
Qy	741	-----	741
Db	161816	TGTTCACTGCTTAGGCGCTCCACCTTAATTGGCAATTACCCATGAATAGTCTGATTTCCA	161757
Qy	741	-----	741
Db	161756	AACTTTGACAGAGGGAAACCTTATGGTTGATCAGTAACCTCCGACCCCAAGTTTGACTG	161697
Qy	741	-----	741
Db	161696	TAAACAAGTTTTTTTAAACAACAATTTGTCTACATCACACTATCATCTCTCAAG	161637
Qy	741	-----	741
Db	161636	TCTAAGAAGATGAGTATCCCTGTGTACCTAACGGGAACCTGAGGCTCAGAGAGGA	161577
Qy	741	-----	741
Db	161576	GGAACCTTGTCCAAGGTACACAGAGCCTAGGCTACAACTGGGACTCTGTGTTTCCAGGCA	161517
Qy	741	-----	741
Db	161516	AAACTCTTCCACTGCACCTACACTGCTGTGGCCCTGCAATGGACACCTCAGTGGCCCA	161457
Qy	741	-----	741
Db	161456	TCATGATACCTCCCTCACCCCTAACTTCCAGCCACATGGTTTAGCAGAGGGAGTAATCAA	161397
Qy	741	-----	741
Db	161396	CCAGAAGCCATGAAGCTTGGGTTCTTTTCTTTTGTGAGACAGAGTCTCACTCTGTG	161337
Qy	741	-----	741
Db	161336	CCAGGCTGGAGTGCATATGATATCTCAGCTCACTGCAACCTCTGCCCTCTGGGTCA	161277
Qy	741	-----	741
Db	161276	AGCAATCTTGTGCTCGGCTCTTGAGCAGCTGGGATACAGGTGTGCACCATACCTGG	161217
Qy	741	-----	741
Db	161216	CTAATTTTGTATTTTAGTAGAGCGGGTTTACCATGTTGCCAGGCTGGCCTTGAA	161157
Qy	741	-----	741
Db	161156	CTCCGACCTCAGGTGATCACCCTCTTGGCCCTCCCAAGTGTGGGATTACAGGTGTG	161097
Qy	741	-----	741
Db	161096	AGCCACCAACCAGCTGATCTTAATTCAGTTTGTTCATTTCTTCATTCAGCAATAT	161037
Qy	741	-----	741
Db	161036	TTGCTGAGTGTCTACTAAGTGCAGCTGCTTTGGGNAATTTAGTGGTGGACACAAAT	160977
Qy	741	-----	741
Db	160976	AAGGTTCTTGCTCTCTGGGGTTTATTTATGTTTCTAATAATAACACTTATAAATATAG	160917
Qy	741	-----	741
Db	160916	TGTTTATTAATTTCAATTTCTTATCTATTTTTTAAATAATAAATAATTTAATCTCTATA	160857
Qy	741	-----	741
Db	160856	ACAACCCCATGAGTAGGTACTATCATTAAGATAGTAATGGAGACACAGAGAAGGTAG	160797
Qy	741	-----	741
Db	160796	TAATTTCTCCAAGGTACACAGCTAGTAAGTGTGAGAGCTGGAATTTGAACCAAGGCAGC	160737
Qy	741	-----	741
Qy	741	-----	741
Db	160736	CTGGTTCCAGAGTCTGGTTCCATGGTTATTCATTTGCTGTGATTTTGTGCTCTCTCTTTA	160677
Qy	741	-----	741
Db	160676	TTACGTTTTCACAGACTAGTGTTCACAGTCAAAATCAAATTTCTTTGGAAGCCTCAATTG	160617
Qy	741	-----	741
Db	160616	CCTTATCTGTGAATAGGAATGATGCCACTCAACTCACAGGATCAAAAGGATGGGTCAAA	160557
Qy	741	-----	741
Db	160556	GAAACAGCCTTGGGAACCTGTAACCACTATACATATTTAAGGATGATATCTCCCTCC	160497
Qy	742	-----LeuProIleGlnArgGlyProProLeuP	751
Db	160496	TTCTCTCACACTCTAGTTTTTACCTTTTACAGCATGCTATCCAGAGGGGACCCCTGCC	160437
Qy	751	oProAlaHisValHisAlaGluSerAenSerThrSerIleTrrLeuArgTrpVal	771
Db	160436	TCAGGCCACGTCATGCGAATCAACAGCTCCACATCATCTGCTTCGGTGGAAAAA	160377
Qy	771	sProAspPheThrValIleValAsnTyrThrValArgPheGlyProTrpGlyLe	791
Db	160376	GCCAGATTTTCCACACAGTCAAGATTGTCAACTACACTGTGTGGCTTCAGCCCTGGGGCT	160317
Qy	791	uArgAsnAlaSerLeuValThrTyrThrSer	802
Db	160316	CAGGAATGCTCCTCTGTCACCTATTACACCAG-GTGGGACTAATTCCTTTACTAATGT	160258
Qy	802	-----	802
Db	160257	AAAGAGGGAGGAAGAAATAGTGGGGGTGGTGGGAAGGACGAGCGTGGTGCATCAGG	160198
Qy	802	-----	802
Db	160197	GAAGGCCCTTGGATTTCAGGAACATGTCTATGGCTCACCTGTATCGCTGAGGCCACACAG	160138
Qy	802	-----	802
Db	160137	AGGGGCTGCACAGGCTCCACGCTCTTGGCTGTGACAGATGTAGCAAGAGGCCATCC	160078
Qy	802	-----	802
Db	160077	AATCCGCTCCCTCAACAACAGAGCCTTGTTCATAACTCCGGAGATGTGACACAGACA	160018
Qy	802	-----	802
Db	160017	GACCTGTTTCTCCAACTCAGAGGGCTCCAGTCTGTATGGTGAAGAGAACCCACCCCTG	159958
Qy	802	-----	802
Db	159957	TCCTCAGAGCTGCTCCCTCTACCTGTAGTAGGGGAGACAGAGCCTGAGGAGCCTGAGG	159898
Qy	802	-----	802
Db	159897	TTTGGCGGAGAAAGTCGCATAAAAGCAACCGATATATACAATGCTGCAATGAAGGCAAAA	159838
Qy	802	-----	802
Db	159837	TTAAGCCAAAGAAAGATGATCGAGGATGATTTCCATAGGTGGAGATTAATCTTGGATTA	159778
Qy	802	-----	802
Db	159777	TCCAGGAAGTTTCTCTGAGGAAGGGGTCTGGTTTCGAGATTGCACTTGGAGCAGCATC	159718
Qy	802	-----	802
Db	159717	ATCCAGACTTCTCTGTAGGTCCCTGCGGTGGTGTAGTGTGCTCCAGCCAGGAGCAGCT	159658

QY	802	-----	802	Db	158579	AGGTATGAGCACAGGCGAGCCAGGTTTCCCTGATGACTTCCAAGCCCCCTGCTGTGG	158520
Db	159657	TCCCTTCCAGGCTAGAGCTGGGGCTGGGCTGACTGCGAGGGAGTAGATTGTGGAGAG	159598	QY	947	-----	947
QY	803	-----	803	Db	158519	ACACAACCTACAATTTCATTGTTTGGACAGCTGCTCTCAGGTTTGGTGGGGGAGAGCA	158460
Db	159597	TCTGCCCTGCTCGGAGAGTGTCTCGAGAGCACTAGTCACGTTTCTGCTCTTAGCAGTTC	159538	QY	947	-----	947
QY	803	rGlyGluAspIleLeuIleGlyLeuIlePheThrLysTyrGluPheAlaValG1	823	Db	158459	AATGTGAAGGGGAATAGGTGTGGTTCAGTGATTAACCTTGATTTCAGAGGCCAGGTTCTGTG	158400
Db	159537	TGAGAGACATCTCTATTTGGGGCTTGAGGCATTCACCAATACGAGTTTGCAGTGCA	159478	QY	947	-----	947
QY	823	nSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSerThrLeuPr	843	Db	158399	GGGACTGGCTTTTTTTTTTTTTTTTTTTCTGAGATGAGAGTTTCACTCTTGTGCCCAGGCTAGA	158340
Db	159477	GTCTCAGCGGTGGACATGATGGGCTTTCCGCTCTGTGGTGGAGCGCTCCACCTGCC	159418	QY	947	-----	947
QY	843	oAspArg	845	Db	158339	GTGCATGGGCGCAATCTCGGCTCACTGCACCTCCGCTCCAGGTTCAAGCGATTCTGC	158280
Db	159417	TGACCGTGAGTGGCTCCAGCCCTTGGCCACCAACCATCTTTTGGCCCACTCTTCTTG	159358	QY	947	-----	947
QY	845	-----	845	Db	158279	TGCCTCAGCCTCCCAATAGCTGGGATTACAGGCACACACCACTGCCCCGGCTAATTTT	158220
Db	159357	CCCCAGCAGCAAGAACAATCTTCAAGACCCCAAGTTTGGGCAGAGCTGGGGGCGAAG	159298	QY	947	-----	947
QY	845	-----	845	Db	158219	TGTATTTTAGTAGAGACAGGATTTACCATGTTGGCCAGGCTGGTCTCAAACTCCTGCAC	158160
Db	159297	AGTTTCTGGGGGAGGAATGTTAAGTCCCTCCAAACACACATACTGTTTCTAGAATT	159238	QY	947	-----	947
QY	845	-----	845	Db	158159	CTCCTGATCTGCTTACCTTGGCCCCCTAAAGTGTGGGATTA CAGGGGTGAGCCACCATG	158100
Db	159237	CCCCAACTAGAAACAGACCTCATCTTTGTGCTTGTGCTTGTTCATGAAGGAGGAAGAA	159178	QY	947	-----	947
QY	845	-----	845	Db	158099	CCCAGCCAAATTATTGTATTTTATTAGTAAGAGGGGTTTTCACCACTGGCCAGGCTGGT	158040
Db	159177	GGGAGGTATATCATCATGCTCTGTCCCCAGCTCTGACATCTCTCCCATCTCTCTCTCC	159118	QY	947	-----	947
QY	846	-----	846	Db	158039	CTCAAACTCCTTACTCAGGTGATCCACCCACCTCGTCTCCCAAAGTGCTGGGATTACA	157980
Db	159117	CGCACCCAGGGCCCTCC-ACACCCCATCCGACCTCGGACTGAGCCCTGCACACCGTC	159059	QY	947	-----	947
QY	861	rThrValArgLeuHisTyrCysProThrGluProAsnGlyGluIleValGluTyrIle	881	Db	157979	GGTGTGAGCCACTGCATCCAGCTGGGACTGGCTTTTACAGACCAGCTTCTTTGGGAAATG	157920
Db	159058	CACGGTTCGGCTGCATCTGTGTGCCCCCACCAGAGCCCAACGGGGAGATCGTGGAGTATCT	158999	QY	947	-----	947
QY	881	uIleLeuTyrSerAsnAsnHisThrGluProGluHisGlnTyrThrLeuLeuThrG1	901	Db	157939	CTTTGATTGGAGGATTAACTTGGTGAAGTGATGAGCTTTATGTGTGTGAGGTGGACAGCA	157860
Db	158998	GATCCTGTACAGCAGCAACACACGAGCCTGAGCACGAGTGACCTTGCTCACCAACGA	158939	QY	947	-----	947
QY	901	u	901	Db	157859	GGACTGGCTACGTAATTTCCAGGGCCAGTGCAGAAATGAAAAAGGTGCTATCCCTTATTTT	157800
Db	158938	-GGGTGAGGGCTCTGCCCCAGGGCATCTGCTCAAGAGCTTCCAGCCTCTCCCGCTTCC	158880	QY	947	-----	947
QY	901	-----	901	Db	157799	AAAAATTAGAGGTCCGACGGCCAGCCGCGGTGGCTTATGCTGTAATCCAGACCTTTGG	157740
Db	158879	AGCAGTGAGAACAGTCTGGGGTTTGTTCAGGGGCTACTGGGAGAGCAAGCAAGCT	158820	QY	947	-----	947
QY	901	-----	901	Db	157739	GAGGCCAAGGTGCGGGCGGATCATGTGTCHAGGAGATCGAGACCATCTCTGGCTAACCG	157680
Db	158819	GCTGCCTAAGTGGAGAGCAGACTCTCCAGTCCCTCTCCCTGCCACCTCTGCATCCACCC	158760	QY	947	-----	947
QY	902	-----	902	Db	157679	TGAAACTCTGTCTCTACTATAAAATACAAAAAAATTAGCTGGGGGTGGTGGTGGATGCC	157620
Db	158759	ACCCCATCCACCATCCCTGCTCTCTCTGCTCTCTCCCTCACCCCGAGGAACATCTTCAGTGC	158700	QY	947	-----	947
QY	907	aGluValHisGlyLeuGluSerAspThrArgTyrPhePheLysMetGlyAlaArgThrG1	927	Db	157619	TGTAGTCCCACTACTCAGGAGGCTGAGCGAGGAAATGGCGTGAACCCCGGAGGAGGAG	157560
Db	158699	TCAGGTCCATGCGCTGGAGAGCGACACTCGGTACTTCTTCAAGATGGGGCGCGACAGA	158640	QY	947	-----	947
QY	927	uValGlyProGlyProPheSerArgLeuGlnAspValIleThrLeuGlnIleuThrPheSe	947	Db	157559	CTTGCAGTGAGCCGAGATCATGCACTCCAGCTCCAGGCTGGGCAACAGAGCGAGACTCCG	157500
Db	158639	GGTGGAGCTGGGCTTTCTCCCGCTGCAGGATGTGATCACGCTCCAGGAGAACTGTC	158580	QY	947	-----	947
QY	947	r	947				


```
Qy 209 TyrArgCysValAlaThrAsnSerAlaArgGlnArgPheSerGlnGluAlaSerLeuThr 228
Db 88616 CAGAGAGTCCAGCAGAGAGAGAGAAATGAGATTCCGACCAGGACGGTTGTTGTTT 88675
Qy 229 ValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAspValValIleValAlaAla 248
Db 88676 ATCCGGGTGCAGAGTGTCTCTGAGGTTCGGAGACTGAAGGAGGAGGAGTCCCTAGCGTC 88735
Qy 249 ProGluAsnThrThrValValSerGlyGlnAsnValValMetGluCysValAlaSerAla 268
Db 88736 CCTCCACG-----GCTCTCCC 88753
Qy 269 AspProThrProPheValSerTrpValArgGlnAspGlyLysProIleSerThrAspVal 288
Db 88754 TCTCTCTCCCA-----GACGGGAAGCCCATCTCCACAGATGC 88792
Qy 289 IleValLeuGlyArgThrAsnLeuLeuIleAlaSerAlaGlnProArgHisSerGlyVal 308
Db 88793 ATCGTCTCTGGCGCCACCAACCTACTAATTGGCAACGCGCAGCCCTGGCACTCCCGCGTC 88852
Qy 309 TyrValCysArgAlaAsnLysProLeuThrArgAspPheAlaThrAlaAlaGluLeu 328
Db 88853 TATGTCTCGCGCCCAACAGCCCGCCACGCGCGACTTCGCCACTGCACGCGCTGAGCTC 88912
Qy 329 ArgVal----- 330
Db 88913 CGTGTCTCTGGTGAAGCGCGCCGAAAGCGCGAGGAGCGGGTTAGCTGGGGCTGGGGTGA 88972
Qy 330----- 330
Db 88973 ATAGAGCAGGGAGGTGACGGGTGGTGTCTGGGGGCACCTCAGCGGCTCTCTCGTAG 89032
Qy 330----- 330
Db 89033 GACGAACTCAGGGAGGAGTCCCTAGTAAGCGCGAGGTGGGACGAGGGATGGGGACTTGGT 89092
Qy 330----- 330
Db 89093 AACTGGAGGACGGAGTGCGGCGGTGTGAGTCTGGGATGGAGCGAGGGGTGATTAAAGC 89152
Qy 330----- 330
Db 89153 CTGGAGCGGTGATTAGCGGGAGTTGTGACTGGGGTTATTCTGGGGATAGTGGAGTGTAT 89212
Qy 330----- 330
Db 89213 CTGAGTGTGTACTAAGAAACTGGGGGCGAGGAGTAGATCTGTTAGGTGTGGACATATGGG 89272
Qy 330----- 330
Db 89273 AGGGCAGGAGGAAGAGATGGGTAAATTTAGGGGTGGGTAGATCCGGGAGGAGCAAT 89332
Qy 330----- 330
Db 89333 ATTTAGGGGAGCTGAGGAGGCTCGGGACAGTTTGGGAGAGAAGAGATGGTCTGTTG 89392
Qy 330----- 330
Db 89393 TGGGGCTGGGGTAAATGGGAGAGAAAGGGTAGGTCTGCGCTGAGCCGAGTAGTTTGG 89452
Qy 330----- 330
Db 89453 AAAGGGTGAATCAGGGAGAGGTTAAAGTGAATTTGAAGGGAGCGAGGGAGGGCGGTGGG 89512
Qy 330----- 330
Db 89513 GGCAGGCTTGGGGTACTGGGGGGGGGGGAGACACTTAGTTGCGGGATCCCGGCTCAGC 89572
Qy 331-----LeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerAr 345
Db 89573 GTCGCCCTCTCGCCCTAGCGCTCCCGCCATCACTCAGGCGCCCGAGGCGCTGTGCGG 89632
Qy 345 gThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLe 365
Db 89633 GACGCGGCGCAGCACAGCGCGCTTCGTGTGCCGCGCTCGGGGAGCGCGCGCAGCGCT 89692
Qy 365 uHisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgValIleValGlnGlyG 385
Db 89693 GCGCTGGCTGCACAAACAGGGGCGCGCTGCGGCCAACGCGCGGTCAAGGTCCAGGGCGG 89752
Qy 385 YGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrTyrGlnCysVa 405
Db 89753 CGTGGCAGCGCTGTTCTATCACAGATCGGCTTCGAGGAGCGCGGCTACTACCAAGTGGT 89812
Qy 405 lAlaGluAsnSerAlaGlyThrAlaCysAlaAlaProLeuAlaValValValArgG 425
Db 89813 GGCTGAGAACAGCGCGGGAATGGCGTGGCTGCGCGCTGCTGGCGTGGTGGTGGCGCA 89872
Qy 425 uGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerValLe 445
Db 89873 GGGGCTGCCAGCGCCCCACGCGGGTCACTGTCTACGCCACTGAGCAGCTCCGCTGTGTT 89932
Qy 445 uValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTy 465
Db 89933 GGTGGCTTGGAGCGGCGCGAGATGACACGAGCAGATCATCGGCTTCTCTCTCCACTA 89992
Qy 465 rGlnLysAlaArgGlyVal----- 471
Db 89993 CCAGAGGACAGCGGGTAC-GTCCGCTGGCGGGGATGCACCAGCGCGCGCGGGCGG 90051
Qy 471----- 471
Db 90052 GCGCAGCCAGAAAGGGTTTCAGTCCGATGACTGAAGGGTTTCCTGTATGGGTACCTCTCG 90111
Qy 471----- 471
Db 90112 GCAAGGATGGACTGGGGTAGGAGCCCATAGGAGAGACCCGCAATAGTATGTTAGTATGG 90171
Qy 471----- 471
Db 90172 ATTAGGCAACTGGCGCTGGAAGGCCACTGCTTCCCTCAGTCATGTTTGTGGAGAGAG 90231
Qy 471----- 471
Db 90232 GGGAGAGGACGTTCTCTGGAGCCACTCCCATCCAGTTAACTAATCTGGAGAGGGGGTCA 90291
Qy 471----- 471
Db 90292 TTTCCAGAAAGCACTCCCTAACCACTAGTGGAGGAATTAATACACTGAGACAGGAGCTGAG 90351
Qy 471----- 471
Db 90352 GAGTTTATACACTCCAGCTCAGATTATCTGGGATGATTTAGCTATATAATAGACACCAG 90411
Qy 471----- 471
Db 90412 CAATATGAAGGATTCCTATACAGACTCCAGCTAGCCAGACAGGGGGGTGGCACTATGAC 90471
Qy 472-----AspAsnValGluTyrGlnPheAlaLys 481
Db 90472 AGTCCCTTGTCCCCCGTGTCCAGGCATGACCAATGTGGAAATACCAGTTTGCAGTGAA 90531
Qy 481 nAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPheTy 501
Db 90532 CAACGACACCAACAGACTACAGGTTTCGGACCTGGACCCCAACACAGATTATGATTTCTA 90591
Qy 501 rValValAlaTyrSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisTh 521
Db 90592 GGTGGTGGCTTACTCCAGCTGGGAGCCAGCGCACCTCCACCCCGCAGCTGGTGGTCACAC 90651
Qy 521 rIleuAspAspVal----- 525
Db 90652 ACTGGATGATGG-TAGGGCCCTCTGAACCTCGCAGTGGGCGAGCTTGGGGTAAAGAGAAATTC 90710
Qy 525----- 525
```


QY	627	-----	627	QY	741	-----	741
Db	92931	TGTTTGACACATTGAAGCTCTCACTCTCAAGTGGGTCTCCCTACTGTATATGGTGGCAT	92990	Db	94009	CCTAGGATCAGGGTCTCTCTTTTAGTCCAGAGAGCCCTGTCTCATGAGAGAGCTTTGTT	94068
QY	627	-----	627	QY	741	-----	741
Db	92991	CCCTCTTCTGGCCTCAGCAACCAACCCCACTGCCCCAGGTGTTGTATCTTCCAT	93050	Db	94069	TGTTCACTGTCTTAGGCTGCCACCTTATTTTGCATTAACCATGAATAGTGTGATTTCCA	94128
QY	627	-----	627	QY	741	-----	741
Db	93051	AGACTTCAAGAGCGCGATTTCTGTCCAGGTGCTATCAGCCTACTCCAGAGGTGCTCTT	93110	Db	94129	AAGTTTGAACAGAGGGAACCCCTATGTTGATCAGTAACCTCCCGACCCCAAGTTTGACTG	94188
QY	627	-----	627	QY	741	-----	741
Db	93111	GCTGACCCCTTTGCAATGCCAGCGGACACCCCTGCATTCCTGCACCTCTGTGTGTTGTA	93170	Db	94189	TAAAAACAAGTTTTTTTTAAAAACACATTTGTCTACATCACACTATCAGACTATCTCTCAAG	94248
QY	627	-----	627	QY	741	-----	741
Db	93171	GAACAGTGAGTTCACTCTAGACCAAGCCAGGACTGGCACTTAGAGTGGTGGCTGAGTG	93230	Db	94249	TCTAAGAAGAGTGAGTATCCCTGTGTGACCTTAACGGGAAACTGAGGCTCAGAGAGGAATA	94308
QY	627	-----	627	QY	741	-----	741
Db	93231	GTGATGTGTGGAGATGCTGGCAAAAGGCTCTCACTGAATGCCATCTTCTGTCTCTT	93290	Db	94309	GGAATTTGTCCAGGTCAACAGGTCAACAGAGCCTAGGCTAACCTGGGACTCTCTGGTTTCCAGGCA	94368
QY	627	-----	627	QY	741	-----	741
Db	93291	CCTCATGCCCTTCTCCCTGTCTCTATFCCCTGGCCTCTTCTCTCATTTCTGCATTTCTGTT	93350	Db	94369	AAACTTTTCCACTGCACACTACACTGCTTGTGGCCCTGCAATTGGACACCTCAGTTGCCCA	94428
QY	628	-----	636	QY	741	-----	741
Db	93351	CCTTAATTTTCCATCTCTCTCTGCCCCCAGTCCCTTTTGGCCCTGCAGAGTTGAA	93410	Db	94429	TCATGATACTCCCTCACCCCTAACTTCCAGCCACATGGTTTACAGAGGGAGTAATCAA	94488
QY	636	sValArgAlaLysMetGluSerLeuValValSerTrpGlnProProHisProThrGln	656	QY	741	-----	741
Db	93411	GGTGAGGCAAGATGAGTCCCTGTGTGTGTCATGGCAGCCACCCCTCACCCACCCCA	93470	Db	94489	CCAGAGCCATGAAGCTTGGGTTCTTTTTTTTCTTTTGAGACAGAGTCTCCTCTGTTG	94548
QY	656	nIleSerGlyTrpLysLeuTrpGlyValGlyThrGluGluAlaAspGlyAs	676	QY	741	-----	741
Db	93471	GATCTCTGCTACAACTATATTGGCGGAGGTGGGGCTGAGGAGGAGGCCAATGGCGA	93530	Db	94549	CCCAGCTGGAGTCAATGGTATGATCTCAGCTCACTGCAACCTCTGCTCTCTGGGTCA	94608
QY	676	pArgProGlyGlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLys	696	QY	741	-----	741
Db	93531	TGCGCTGCACAGGGGCCGTGGAGACAGGCTTGGGATGTGGGCTGTCCGCTCAAGAA	93590	Db	94609	AGCAATTTCTGTGCTCGGCTCTTGAGCAGCTGGGATTACAGGTGTGCACCACCTACCTGG	94668
QY	696	sLysValLysGlnTrpGluThrGlnLeu	706	QY	741	-----	741
Db	93591	GAAGTGAAGCAGTATGAGCTGACCCAGCT-AGGTGAGGAGGGCTGGGGGAGGGGAGG	93649	Db	94669	CTAATTTTGTATTTTGTAGTAGAGACGGGGTTTACCATGTTGGCCAGGCTGGCCTGAA	94728
QY	706	-----	706	QY	741	-----	741
Db	93650	GAGGGATGAGCTTTCAGTGGGAGCTGGGCAAGTTGTGTGAAGAAGTTAGAGATGGGATGC	93709	Db	94729	CTCCCGACCTCAGGTGATCCACCCATCTTGGCCTCCCAACGTGCTGGGATTACAGGTGTG	94788
QY	706	-----	706	QY	741	-----	741
Db	93710	TCTCTGTGAGCGGAGAGTGACCCCAAGATGTCCCCCTAGCCATTAGCCCTGTACCTTC	93769	Db	94789	AGCCACCACACCCAGCTGAGTTCTAAATTCAGTTTGTCTCAITTTCTTTCAGCAAAATAT	94848
QY	707	-----	725	QY	741	-----	741
Db	93770	CCAGTCTCTGGCGGCTGTACGAGGTGAAGCTCGTGGCTTTTCAACAAACATGAGGATGG	93829	Db	94849	TTCCGTGAGTGTCTACTAAGTGCAGCTGCTTTTGGGAATTTAGTGGTGCAACAACAAT	94908
QY	725	YTrpAlaValTrpGlyThrGluLysAlaProThrProAsp	741	QY	741	-----	741
Db	93830	CTATGACAGAGTGTGAAGGGCAAGCAGAGAGGCGCCGCCACCCAG- TGAGGGCGGTG	93888	Db	94909	AAGTTCTTGTCTCTTGGGGTTATTTTATGTTTCTAAATAAACACTTATAAATATAG	94968
QY	741	-----	741	QY	741	-----	741
Db	93889	GGGGAAGAAGGGGGAGGGCTCAGGTGAGGGCTCTGCTAGCGGCTTCTCTCTTTTCC	93948	Db	94969	TGTTTATTATTTCATTTTCTATTTTCTATTTTAAAAATTTAAATAATTTAATCCTCATA	95028
QY	741	-----	741	QY	741	-----	741
Db	93949	TTTGGGCCAGCTTGGTGAGCAGGTACTTTTAGTAGCCTCAGCTCATGCCCAACCCATCTT	94008	Db	95029	ACAACCCCATGAGGTAGGTACTTATCATTTAGATAAGTGAATGGAGACACAGAGAGGTAA	95088

Db	95089	TAATTTCTCAAGGTCA	CACAGCTAGTAAGTGT	CAGAGCTGGAATTTGA	ACCAAGGCAGC	95148
Qy	741	-----	-----	-----	-----	741
Db	95149	CTGTTCCAGAGTCTGGT	TCATGGTTATTCATTG	CTGTGCTCTCTCTTTA		95208
Qy	741	-----	-----	-----	-----	741
Db	95209	TTACGTTCTTACAGACT	AGTGTTCACAGTCAMAA	ATCTTGGAGCCTCAATTG		95268
Qy	741	-----	-----	-----	-----	741
Db	95269	CCTTATCTGTGAAGT	AGATGATGCCACTCA	ACTCAGAGTCAAAAGG	TGGGTCAAA	95328
Qy	741	-----	-----	-----	-----	741
Db	95329	GAACAGCCTTGGAACT	GTATAACCACTATACA	TATTTAAGGATGATAT	CCTCCCTCCC	95388
Qy	742	-----	-----	-----	-----	751
Db	95389	TTCTCTACACTCTAGT	TTTTCACCTTTCAGAC	ATGCTTATCCAGAGGG	ACACCCCTGCC	95448
Qy	751	oProAlaHisValHis	AlaGluSerAsnSerSer	ThrSerIleThrLeuArg	TrpIysLy	771
Db	95449	TCCAGCCCCAGTCCAT	CGGAATCAAAACAGCT	CCACATCCATCTGGCT	TCGGTGGAAAA	95508
Qy	771	sProAspPheThrThr	ValValIleValAsnTy	rThrValArgPheGly	ProTrpGlyLe	791
Db	95509	GCCAGATTTACCACT	GTCAAGATTGTCAACT	ACACTGTGCGCTTCAG	CCCCCTGGGGCT	95568
Qy	791	uArgAsnAlaSerLeu	ValThrTyrrTyrrThr	Ser-		802
Db	95569	CAGGAATGCCCTCCCT	GTGTCACCTATTACAC	CAG-GTGGGGACTAAT	GCTTACTAATGT	95627
Qy	802	-----	-----	-----	-----	802
Db	95628	AAAGAGGAGGAAGAAT	AGTGGGGGTGTGTGG	GAAGGCAGCAGGTGT	GTGATCAGG	95687
Qy	802	-----	-----	-----	-----	802
Db	95688	GAAGGGCCCTGGGATT	CAGGAACATGTCTATG	GTCTACCTGTTACGT	TACGTGAGGCCAC	95747
Qy	802	-----	-----	-----	-----	802
Db	95748	AGGGGTGCACAGGGT	CCACGTCTTGCTGCT	GACCAAGATGTAGCA	AGGAGGCCATCC	95807
Qy	802	-----	-----	-----	-----	802
Db	95808	AATCCCGTCCCTCACA	ACAGAGCCTTGTCTCA	AAACTCCGGAGAATGT	GCACACAGGACA	95867
Qy	802	-----	-----	-----	-----	802
Db	95868	GACCTGTTTCTCAACT	CTCAGAGGCTCCAGT	CTGTGATGTGAAGAA	CCCCCACCCTG	95927
Qy	802	-----	-----	-----	-----	802
Db	95928	TCCTCAGAGCTGCTCC	CTTACCTCTTAGATG	GGGGAGACAGACCTG	AGGAGCCTGAGG	95987
Qy	802	-----	-----	-----	-----	802
Db	95988	TTTGGCGGAGGAAGT	CGCATATAAAAGCAAC	CCGATATATACAACT	GCCATGAAGGCAAAA	96047
Qy	802	-----	-----	-----	-----	802
Db	96048	TTAAGCCAAAGAAAGAT	GCATCGGAGATGATTC	ATAGTGGAGAAATAACT	TTGGATTAA	96107
Qy	802	-----	-----	-----	-----	802
Db	96108	TCCAGGAAGGTTTCTC	TGGAGGAAGGGGTCTG	GTTCGAGATTGACATT	GGAGCGACATAC	96167
Qy	802	-----	-----	-----	-----	802
Db	96168	ATCCAGACTTCTCTGT	AGGTCCCTGGTGTGTAG	TCTCGCTCCAGGCCAG	GGAGCAGCT	96227
Qy	802	-----	-----	-----	-----	802
Db	96228	TCCCTTCCAGGGTGTAG	AGCTGGGGCTGGGGCTG	ACTGACAGGGAGTAG	ATTGGAGGAG	96287
Qy	803	-----	-----	-----	-----	803
Db	96288	TCTGCTCTGTGGAGAG	TGTGCTGAGAGCACT	AGTACAGTTTCTTGT	CTTAGCAGTTT	96347
Qy	803	-----	-----	-----	-----	803
Db	803	rGlyGluAspIleLeu	IleGlyGlyLeuLys	ProPheThrLysTy	rGlyPheAlaVal	823
Db	96348	TGAGAAGACATCTCAT	TGTGGCGCTTGAAGC	CAATTCACCAATAG	AGTTTGCAGTGCA	96407
Qy	823	nSerHisGlyValAsp	MetAspGlyProPhe	GlySerValValGlu	ArgSerThrLeuPr	843
Db	96408	GTCTCAGCGGTGGACA	TGATGGGCTTTTGG	CTCTGTGTGTGAG	GGCTCCACCTGCC	96467
Qy	843	oAspArg	-----	-----	-----	845
Db	96468	TGACCGTGAAGTGTCT	CCAGCCCTTGGCCAC	CAACCCATCTTTTGT	GCCAATCTCTCTTG	96527
Qy	845	-----	-----	-----	-----	845
Db	96528	CCCCAGCAGCAAGAA	ACATCTTCAGAACCC	CAAGGTTTGGGCAG	AGCTGGGGGCAAG	96587
Qy	845	-----	-----	-----	-----	845
Db	96588	AGTTTCTGGGAGGGG	AGGATGTTAAGTCC	CTCCACACACATAC	TATGTTTTCTAGAA	96647
Qy	845	-----	-----	-----	-----	845
Db	96648	CCCCAACTAGAACCA	CAGACCTCATCTTGT	CTTGTGTGTTCAT	GAGGAGGAAGAA	96707
Qy	845	-----	-----	-----	-----	845
Db	96708	GGGAGTCAATCATCAT	CTCTGTCTCCCGAG	CTCTGACATCTCTCC	CAATCTCTCTCTCC	96767
Qy	846	-----	-----	-----	-----	861
Db	96768	CGCACCCCGGGCCCT	CTCC-ACACCCCAT	CTCGACCTGCGACTG	AGCCCCCTGACAC	96826
Qy	861	rThrValArgLeuHis	TrpCysProThrGlu	ProAsnGlyGluIle	ValGluTyLe	881
Db	96827	CACGGTTGGCTGCAT	CTGTGCCCCCAGAG	CCCAACAGGGAGAT	CTGTGAGTATCT	96886
Qy	881	uIleLeuTySerAsn	AsnHisThrGlnPro	GluHisGlnTrpThr	IleuThrThrG1	901
Db	96887	GATCTGTACAGCAG	CAACACACGACGCT	GTAGCACCACTTGT	CTCACCACGCA	96946
Qy	901	u	-----	-----	-----	901
Db	96947	GGGTGAGGGCTCTCT	CGCCCCAGGGCATCT	GTCTCAAGAGCTT	CCAGCTCTCTCGGCT	97005
Qy	901	-----	-----	-----	-----	901
Db	97006	AGCAGTGAGAAACAG	TGCTGGGGTTGTTT	CCAGGGGCTACTGG	GAGAGCAAGCAG	97065
Qy	901	-----	-----	-----	-----	901
Db	97066	GCTGCCCTAAGTGG	GAGACAGACTCTCC	AGTCCCTCTGCCCC	CACTCTGCAATCC	97125
Qy	902	-----	-----	-----	-----	907
Db	97126	ACCCCATCCACCAT	CTCCCTGCTCTCC	CTGCTCCACCCCA	GGAAACATCTTCA	97185
Qy	907	AGluValHisGlyLeu	GluSerAspThrArg	TyrPhePheLysMet	GlyAlaArgThrG1	927
Db	97186	TGAGTCCATGGCTTG	GAGAGCGACACTCG	TACTTCTTCAAGAT	GGGGGCGGCAC	97245
Qy	927	uValGlyProGlyPro	PheSerArgLeuGln	AspValIleThrLeu	GlnGluThrPhe	947
Db	97246	GGTGGGACCTGGGCT	TTTCTCCCGCTTGC	AGGATGTGATC	ACGCTCCAGGA	97305

Qy	1230	LysAlaLeuCypProLeuThrValSerProSerIeuProArgAlaProValSerSerAla	1249
Db	101682	AAATCCCTGCGCCTTAGAGCCACGCCAGCGCTGCCAGATCCCCGGTCTCTCTCTCT	101741
RESULT 15	AC112162/c		
LOCUS	AC112162.8	GI:21327375	
DEFINITION	Mus musculus chromosome 9 BAC clone MGS1-296M6 ES cell line, complete sequence.	DNA linear	ROD 21-SEP-2002
ACCESSION	AC112162		
VERSION	HTG.		
KEYWORDS	Mus musculus (house mouse)		
SOURCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
ORGANISM	1 (bases 1 to 116580)		
REFERENCE	Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zencheck,W., Xi,C., Juels,P. and Kucherlapati,R. High Throughput Mouse Sequencing Unpublished		
TITLE	2 (bases 1 to 116580)		
JOURNAL	Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zencheck,W., Xi,C., Juels,P. and Kucherlapati,R. Direct Submission		
AUTHORS	Submitted (20-FEB-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA		
TITLE	3 (bases 1 to 116580)		
JOURNAL	Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zencheck,W., Xi,C., Juels,P. and Kucherlapati,R. Direct Submission		
AUTHORS	Submitted (04-MAY-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA		
TITLE	4 (bases 1 to 116580)		
JOURNAL	Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zencheck,W., Xi,C., Juels,P. and Kucherlapati,R. Direct Submission		
AUTHORS	Submitted (04-JUN-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA		
TITLE	5 (bases 1 to 116580)		
JOURNAL	Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zencheck,W., Xi,C., Juels,P. and Kucherlapati,R. Direct Submission		
AUTHORS	Submitted (06-JUN-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA		
TITLE	6 (bases 1 to 116580)		
JOURNAL	Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zencheck,W., Xi,C., Juels,P. and Kucherlapati,R. Direct Submission		
AUTHORS	Submitted (21-SEP-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA		
TITLE	On Jun 6, 2002 this sequence version replaced gi:21321767.		
JOURNAL	-----Genome Center: Center: Harvard Partners Genome Center Center Code: HPGC Web site: http://www.hpcgg.org/sequence/mouse.html Contact: hpgc@mendel.mgh.harvard.edu		
COMMENT			

CLONE LENGTH: This sequence represents the entire insert of this clone unless otherwise noted. If there are overlapping clones, the overlaps are noted in the beginning and end of the Features

listing.

ANNOTATION OF FEATURES:

SPNs are identified using ePCR (Genome Res. 7:541-550). Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences. Genes and Regions of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to EST and cDNA sequences in Unigene. Genes demonstrate at least two exons flanked by consensus splice sites that maintain sequence continuity across the consensus splice sites that are not identical matches are annotated as similar.

SEOURNCING READ COVERAGE: Attempts are made to complete double

strand sequence for all regions. All sequence is completed to a standard of coverage with a minimum of 3 reads with no ambiguities. If the sequence coverage for a region does not meet this standard, it is indicated in the annotation as low coverage. Low coverage linkages are verified by PCR product size verification or verification of forward and reverse reads from clones which span the low coverage area.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality

standards - estimated average error rate is less than 1 per 10,000 bases using the Consed quality parameters. Regions that do not meet this requirement are annotated as Low Quality.

-----Summary Statistics

Center project name: ALI
Sequencing vector: pSMART; AP399742
Chemistry: Dye-terminator Big Dye; 100#
Assembly program: Phrap version 0.990319
Contig length: 116580
Fraction of Phrap value < 40: 0
Error Rate in Consed: 0.00 per 10,000 bp
Number of N's in consensus: 0

----- Distribution of Quality < 40 Bases:

A blank coordinate grid for plotting. The vertical axis (y-axis) is on the left, with labels from 0 to 1000 in increments of 100. The horizontal axis (x-axis) is at the bottom, with labels from 5 to 40 in increments of 5. The grid consists of a series of horizontal lines corresponding to the y-axis labels and vertical lines corresponding to the x-axis labels. The origin (0,0) is at the bottom-left corner.

FRACTURES

```

1. .116580
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="ES cell line"
/db_xref="taxon:10090"
/chromosome="9"
/clonome="MG81-296M6"
STS      235. 369
/standard_name="DSM1316"
repeat_region 258. 303
/rpt_family="(TG)n"
STS      612. 1204
/standard_name="Nope"
repeat_region complement(1584..1643)
/rpt_family="B4A"
repeat_region 1710. 1748
/rpt_family="(TG)n"
repeat_region complement(3793..3909)

```


repeat_region	/rpt_family="PB1D10"	4042..4191	Db	12486	GGGTCCTTTGGAGGCTACCGGGGGGAGATGTGGTCAATTTGGCAGCCCCCAGAGAACACC	12427
repeat_region	/rpt_family="MIR"	4247..4435	Qy	253	ThrValValSerGlyGlnAenValMetGluCysValAlaSerAlaAspProThrPro	272
repeat_region	/rpt_family="B2_Mm2"	complement(4504..4623)	Db	12426	ACGGTAGTGTCTGGACAGAGGTAGTGTGATGGAGTGGTGGCTCTGCTGACCCACCCCT	12367
repeat_region	/rpt_family="RSINE1"	4720..4831	Qy	273	PheValSerTrpValArgGln	279
repeat_region	/rpt_family="L2"	complement(4840..4879)	Db	12366	TTTGTGTCTGGGTCCGACAGAGGTGAGTGTGAGAAAGAGGAGGAGAGCTGGGGAG	12307
repeat_region	/rpt_family="MIR"	4938..5118	Qy	279	-----	279
repeat_region	/rpt_family="B2_Mm2"	6216..6250	Db	12306	TGGGCAAGATGGTACAGTATTTTCAGAGCTCCAAAGTCAGAGGCTTCTGCTTAAAGATAGA	12247
repeat_region	/rpt_family="AT_rich"	complement(6251..6381)	Qy	279	-----	279
repeat_region	/rpt_family="RSINE1"	complement(6421..6566)	Db	12246	AGGAGAGATTGACTCTTAGAGTTGCTTTTGACTTCCATAGCACATGCAAGCAAGCTC	12187
repeat_region	/rpt_family="B1_MM"	6593..6612	Qy	279	-----	279
repeat_region	/rpt_family="(T)n"	complement(6668..6790)	Db	12186	ACATGTACACACATTCATGCAGTAATAAAGTCTAAAGAGGTTGTATTAAAGCCAGTGG	12127
repeat_region	/rpt_family="B1_MM"	complement(6705..6820)	Qy	279	-----	279
repeat_region	/rpt_family="MIR2B"	complement(6883..6954)	Db	12126	TGGTGGCACACATCTCTTAATCCAGCATCTGGGAGGCGAGGCGGGCTCTGAGTT	12067
repeat_region	/rpt_family="ID_B1"	complement(6982..7058)	Qy	279	-----	279
repeat_region	/rpt_family="ID4"	8050..8246	Db	12066	TGAGGCCAGCCTGGTCTACAGAGCAAGTTTCAGGGTAAACAGGGGCTACTCAGAGGAAACCC	12007
repeat_region	/rpt_family="B2_Mm2"	complement(8445..8719)	Qy	279	-----	279
repeat_region	/rpt_family="B4"	complement(9019..9119)	Db	12006	TACTTTGGAAAATTCTGGGGTGGGAGGAGAGAGACACAGAGAGACAGAGAGCGAG	11947
repeat_region	/rpt_family="MIR"	9285..9371	Qy	279	-----	279
repeat_region	/rpt_family="MIR"	complement(10773..10846)	Db	11946	GGACAGAGACAGACAGACAGGCGAGACACACACACACACACACACACACACAC	11887
repeat_region	/rpt_family="PB1D10"	10859..10956	Qy	279	-----	279
repeat_region	/rpt_family="B4"	complement(11043..11141)	Db	11886	ACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTGTATTAAATATCTGCTCCG	11827
repeat_region	/rpt_family="PB1D7"	11848..11884	Qy	279	-----	279
repeat_region	/rpt_family="(TC)n"	11884..11922	Db	11826	CCCACCAGAGACCATTTAGCCATCGTTTACCACCTCCCTTCTGCTTTTCCACCTAGGA	11767
repeat_region	/rpt_family="(TG)n"	11930..11966	Qy	279	-----	279
repeat_region	/rpt_family="(TC)n"	complement(11967..11994)	Db	11766	TCAGAGTAGAGAGGAACTGGGATTCAAACCGGGCTGTGATGCAGCTTTTCAGGGATGT	11707
repeat_region	/rpt_family="B4"	complement(12127)	Qy	279	-----	279
repeat_region	/rpt_family="B1_MM"	complement(12128..12143)	Db	11706	AGTGGGAAACCTCGTAGTTGGAGAGGGGTGTAGAGGACAGTTCCAGGCTCCCTCCAGCC	11647
repeat_region	/rpt_family="B4"	complement(12170..12252)	Qy	280	-----	293
repeat_region	/rpt_family="RSINE1"	12853..12993	Db	11646	TGCTCCTTCTCTGCTACCAGATCGAAAGCTATCTCCAGGATGTCTATCGTTCTGGGCG	11587
repeat_region	/rpt_family="B1_MM"	complement(13288..13367)	Qy	293	gThrAsnLeuLeuIleAlaSerAlaGlnProArgHisSerGlyValTrpValCysArgAl	313
repeat_region	/rpt_family="MIR"		Db	11586	GACCAATCTACTCATCGCAGCGCGAGCTCGGCACCTCTGGAGTCTATGCTGCGGAGC	11527
repeat_region			Qy	313	AsnLysProLeuThrArgAspPheAlaThrAlaAlaGluLeuArgValLeu	331
repeat_region			Db	11526	CAACAAGCCCCGACCGGTGATTTTCGCCACTGCGGCTGCTGAGTCCGAGTGTCTGTAA	11467
repeat_region			Qy	331	-----	331
repeat_region			Db	11466	GGAAGGCTGGTCTGGGGGACGTTAAGGATGGGTGTGTGTAGAGGAGAGAGGGCGGT	11407
repeat_region			Qy	331	-----	331

Alignment Scores:

Pred. No.:	2,84e-90	Length:	116580
Score:	2466.50	Matches:	837
Percent Similarity:	20.68%	Conservative:	19
Best Local Similarity:	20.22%	Mismatches:	41
Query Match:	37.35%	Indels:	3248
DB:	10	Gaps:	20

US-09-754-997A-2 (1-1252) x AC112162 (1-116580)

Qy 233 GlySerLeuGluAlaThrArgGlyGlnAspValValIleValAlaProGluAsnThr 252

Db	11406	CCTAGAAGACTAAATGTGTGAGGGGGACCATGAGGAGGTCTGAAATTTTCTGATGGTGG	11347
Qy	331	-----	331
Db	11346	TGATTTGTTGGAGACTGGATTCCTGAGTGTGGGGGGGGGGAGGGAATGGGGCGATTT	11287
Qy	331	-----	331
Db	11286	GAGACGGGGGTGGGGGGCGGATTTAAAGCCTGGTTTGGGTTCATGAGGAGTGAGAGTGC	11227
Qy	331	-----	331
Db	11226	AGCGGAGCAGAGTGGGGTGTAGGTTATGGAAGGAGATAGTGAATGGTCTAGGGGGTGT	11167
Qy	331	-----	331
Db	11166	GTGAAGCCAGAAATCTGTAGTGGTAGCACTCAGAAAACGTGGCAGGAGCGTAGTTGG	11107
Qy	331	-----	331
Db	11106	ATACCAGCCCTAGCTACAGAGTTGAGACCAGCTACAGAGTGAAAGTCTATCTCAAAAAC	11047
Qy	331	-----	331
Db	11046	AAAAATTTAAACAATGACAAATCTCTAGGCTATTAAAGTCCCTATGAGTGTGTGCAGGCC	10987
Qy	331	-----	331
Db	10986	ATGGTGAAGTATTCGGGTCA CAGACTGAGTTGAGACAAGGTTTCTGTTCACCTTCATAC	10927
Qy	331	-----	331
Db	10926	TCCAGCTAGCTGGCCCGGGAAATCTGGGGAATCTCCTGTCTCTCTCTCTAGGATGC	10867
Qy	331	-----	331
Db	10866	TGGGATTAATCAGTTGGGTGTATGTGAGCGCGGAGATCAGAAATTCAGGCGCAACCTG	10807
Qy	331	-----	331
Db	10806	GGCTATAAAGAGCTTTCTCAATACAAGAAAACTGGGGGACCGGCTTGTCTGAAGCT	10747
Qy	331	-----	331
Db	10746	GGAGTAACTGCTGGGTATAGAAGCCACGGGGAGGAAGTCAAGTTACAGCCTAGAGC	10687
Qy	331	-----	331
Db	10686	TCAAAAGCTGGAATCTGTGGCTTGGGAGAAAGGAGGACTAGCCAGGAGGGTCCGAG	10627
Qy	331	-----	331
Db	10626	TCTAGTGGACAGAGCTGGGCCGTATGACTAGGGGAGAGGACCTCAGGGTGAGGCTTGG	10567
Qy	331	-----	331
Db	10566	GCGCGGTGTTGGACACTTGGAGAGGGGTAGGCCCTCTTGTGACAGTCTCCCTCTGTC	10507
Qy	332	-----AlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerArgThrArgAlaSerTh	350
Db	10506	CTCAGCTGCCCCAGCCATCTCGCAGGGCGCCGAGGGCTCTCGCGACCGGGGCCAGCAC	10447
Qy	350	rAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLeuHis	370
Db	10446	CGCGCCTCTGTGTGCGGGCGTCCGGGGAGCCACGGCCCGCGCTGCATGGCTGCACGA	10387
Qy	370	pGlyIleProLeuArgProAsnGlyArgVallyValGlnGlyGlySerLeuVa	390
Db	10386	CGGGATCCCGTTCCGACCCATGGGCGCGTCAAGGTGCAAGGCGGTGGCGGAGCTTGGT	10327
Qy	390	lIleThrGlnIleGlyLeuGlnAspAlaGlyTyrTyrGlnCysValAlaGluAsnSerAl	410
Db	10326	CATCACTCAGATCGGCTTCAGGACGCTGGCTACTACAGTGTGCTAGCAGAAAAACAGCGC	10267
Qy	410	aglyThrAlaCysAlaAlaProLeuAlaValValArgGluGlyLeuProSerAl	430
Db	10266	GGGAACCTGCTGTGCGCTGCGCCCTGGCGGTAGTGTGTGCGAGGGGCTGCCAGCGC	10207
Qy	430	aProThrArgValThrAlaThrProLeuSerSerSerValLeuValAlaTTPGluAr	450
Db	10206	CCGACTCGGGTCA CAGCCACCGCTGAGCAGCTCTCTGTGTGTGGTGGCTGGAGCG	10147
Qy	450	gProGluLeuHisSerGluGlnIleLeGlyPheSerLeuHisTyrGlnLysAlaArgG1	470
Db	10146	GCCTGAGTTGCACAGCGAGCAATCATTGGCTTCTCTCTTCACTACCAAAAAGGAGGGG	10087
Qy	470	yVal-----	471
Db	10086	TAT-GTCTACTCTTGACCGCGGAATCCAGTTAGGCAAGTATGGGATGATAATAA	10028
Qy	471	-----	471
Db	10027	GATGAGAGGAAGGATGGGAGCAATACTGCTGGGAAGTGTATCTAGTATTAGAGATGGA	9968
Qy	471	-----	471
Db	9967	GTGTACAGCACTCCTTGACCGAGAGAAATGCTTCTCTTCCAGGGTTAAAGAGCCATCT	9908
Qy	471	-----	471
Db	9907	TCCACCCCGCACTAACTGACGAGCGGCTCTTCAATTAGCACGGGAGCTTCTCTG	9848
Qy	471	-----	471
Db	9847	GCCTATTATGGAATCAATGCAATAGTAGACTAGATGCGAGAGGCAATTTCTATCTCTCAC	9788
Qy	471	-----	471
Db	9787	CCTAGATTGTCTGTAACTTTAGCTATACATAGACCAGTAGGATCCCGACAGACTCTCC	9728
Qy	471	-----	471
Db	9727	AGCTAGTCAGACATGGAGTGGCTGTGACAGTCCCTTTGTCCCCCGGTGTACCACCTC	9668
Qy	472	-----AspAsnValGluTyrGlnPheAlaValAsnAsnAspThrThrGluLeuGlnVa	489
Db	9667	AGGAGTGGACAATGTGGAGTACCAGTTTGCAGTAAACAATGACACACAGAGCTGCAGT	9608
Qy	489	lArgAspLeuGluProAsnThrAspTyrGluPheTyrValValAlaTyrSerGlnLeuG1	509
Db	9607	TCGGGACCTGGAAACCAACACGGATTATGAGTTCTACGTGTGGCTTACTCCAGCTGGG	9548
Qy	509	yAlaSerArgThrSerSerProAlaLeuValHisThrLeuAspAsp-----	524
Db	9547	GGCCAGCCGAACTCCAGCCCGCCCTGTGTGCATACACTGGACGATGGTAGGACCTCTGA	9488
Qy	524	-----	524
Db	9487	ACCTGACGTGACCTGGGACCCAGAGACACTCCCACTCAGGGGACAGTGTGGCCCAAGGCC	9428
Qy	524	-----	524
Db	9427	TCTTCCCTGCTGTGTGTGTCTGTGTGCCCTATAAAGCAGTGGGCGAGAGGTTCTTCTTC	9368
Qy	524	-----	524
Db	9367	CCAGTGTAAATGGGGAACCTGGGGCTCAGAAAATGAAAATGACTTGGCCAGCTCAGCAG	9308
Qy	524	-----	524
Db	9307	GGCAGCTGGGACTTGAGCCCGAGCTCTGAGCCCCCATACCATGACGCCCTTCTTCTTAACA	9248
Qy	524	-----	524
Db	9247	TACATTTACTACCAGCCATTGGATATGAAAGCTTTCTGGCCCCACACTGTCTCCTAGCTTT	9188

QY 741 ----- 741
Db ACCTAAGTATCTACCATGACAGCTCTGTTCTGGAACTTAGTAATGAACAACATGAAT 4749
QY 741 ----- 741
Db AAGGTATCTGCTCTCTGGAGCTTATTTTATGCTTATCTAATAAACAATTTGTCATCA 4689
QY 741 ----- 741
Db AATAGTATACACTGTCCACAGCTTCTACCTGCTTTGTAAGCAGCCTCCTAAGGACACT 4629
QY 741 ----- 741
Db CATGAGGGGACAGTGAGATGGTTTCAGCAGGGAAGTGATGCTACAGGCCACATGACCTG 4569
QY 741 ----- 741
Db AGCTTGGTTTCTGCACCCCTCCCAAAATGTCCTGCTCTGCTCCCAACATGACACAGCAG 4509
QY 741 ----- 741
Db CACACTCCCTCCCTCCACACATAAATGTGTGACTGAGAGATGACAGGATTAAG 4449
QY 741 ----- 741
Db AATTTAAATATGTTTAAAGATGTATTTATGATTTATATGTTGAGTACACTGTCACT 4389
QY 741 ----- 741
Db GTTGCTATCTTCAGACACAGCAGAGAGGTCCTAGATCCATATATAGATGGTTGTGAGC 4329
QY 741 ----- 741
Db CACCATGTGGTTGCTGGGAATTTGAACTCAGGGACTCTGGAAGAGCAGTCAGTGTAAACC 4269
QY 741 ----- 741
Db GCTGAGCCCTCTCTTCAGCCCTCCAGAAATTTTAAATATTTTAAAGAAATTAATAAAGAA 4209
QY 741 ----- 741
Db AAGAAGAAATGTCAAGATGATGATATGTCATCAGATCCTTATAAGAAATGAGAAATGA 4149
QY 741 ----- 741
Db TGGCACAGAGAGTTAAACAATTTCTTAAGCCACATACATAGTAGGAAGTGGCAGCA 4089
QY 741 ----- 741
Db ATGGAGCTGGCTCCAGAGTCAGGGCATGCTTTCTGTGTACATTTCTCATCAGCCATGGG 4029
QY 741 ----- 741
Db ATATGTACAGTCACTCTATGTGTGACCTTAAGCAAAAAAATCACTTCCATTTCTGGGA 3969
QY 741 ----- 741
Db GCCTCAGTGCAGTAGGATTAATGTACCCAACTCAACAGGATCAAGAGAGGCCCGGTG 3909
QY 741 ----- 741
Db CCAGCTGTGGTATCACACAGCTGTAAACCCCTGTGTAAAGGACCAGAGGCACAGGATGG 3849
QY 741 ----- 741
Db AGAATTACAGGCCAGCTAGGCTACTACATAATGAGATTTCTGCTCAAAAAATAGAGGGG 3789
QY 741 ----- 741
Db GAAAGGCCAATGAACCTACTGAGACCATGTAGTCACCTACTCTCTGACACTGTTGGG 3729
QY 742 -----LeuProIleGlnArglyProProLeuProProAlaHisValHis 756

Db TTTCCCTCTTAGACCTGCTATCCAGAGGGGGCACGGCTGCTCTCTGCCCATGTCAC 3659
QY 757 AlaGluSerAsnSerThrSerIleTTPLeuArgTTPlySlySProAspPheThrThr 776
Db GCAGAGTCAACAGCTCCACTTCCATTTGGCTTCGGTGGAGAGAGCCAGACTTTTACCAC 3609
QY 777 ValIysIleValAsnTyrThrValAtgPheGlyProTTPGlyLeuArgAsnAlaSerLeu 796
Db GTCAGAGTTGTCACATACACTGTAGGCTTCGGCCCTTCGGGGCTTCAGGAATGCTTCCCTG 3549
QY 797 ValThrTyrThr----- 801
Db GTCACCTACTATACACAGGTATCAGTGAGGGGGAGTGTAGCTGGGGTGGTAGACTAGCGG 3489
QY 801 ----- 801
Db GTTCAGCTGTGTGGTAGGCCACACCCAGGATCTACACAGATGTCTTTTCCGTGATGCTGA 3429
QY 801 ----- 801
Db CCTGGGGGTAAATATGAGGGATGGGATTCCTCCCATCCCTCAGTGGGAGAAATGTGTGTC 3369
QY 801 ----- 801
Db AGCTTTTGTGTGAGAGTGGATGCAAGACTAGAGGCTTCATCCGTTGGCTGGCACCTGCC 3309
QY 801 ----- 801
Db GGCCCATCATTCGGAAGATAGCTTAAAGAGCTGCTGTCTGATTTGGGAAGACACCTGTG 3249
QY 801 ----- 801
Db AAAGCCAAACAGAGAGAGAAACTAAGCTGAGGATAGACACTTCAGGAGCGATTTGTCCTCC 3189
QY 801 ----- 801
Db AGGTGCAGAAATTACTTAGATTAATTCGGGAAAGCGTCCAGAGAGAGAGAGGGCTTGCTT 3129
QY 801 ----- 801
Db TGGAGATGAGACATGGAATAACATATGTCGGGCTTTCTGTGTGTTTCCCTGGCTGTGTGT 3069
QY 801 ----- 801
Db GGGAGGGAGGTGTGTGGGCCCCAACAGCAGCTCCCTCCAGGGTGGGGCTGGCTGGAGGC 3009
QY 801 ----- 801
Db AGGGGCTTCTGTCTGGGTTTCTGCTGGCTTGAGGCCCTGTCTGAGGGCAGCTAGTCTTATC 2949
QY 802 -----SerSerGlyGluAspIleLeuIleGlyLeuLysProPheThrIys 817
Db CTTTCTCTGGAAGCTCTGGAGAGACATTTCTCATTTGGCGGCTGAAACCATTTTACCAG 2889
QY 818 TyrGluPheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValVal 837
Db TAGAGTTTGGGTACAGTCCACGGAGTGATATGATGGCCCTTTGGCTTCGTCGTA 2829
QY 838 GluArgSerThrLeuProAspArg----- 845
Db GAAAGCTCCACCTCCCTGACCG- TGAGTATCCACTGCCCCGCTCCCACTGTGGCCACCA 2770
QY 845 ----- 845
Db TTTTGTGGGTCTGAGGAAGGAAGAGAGAGGCTGTGCTCTCATCTCCCTTAGTTAGG 2710
QY 846 -----ProSerThrProProSer 851
Db GCTGTGCTGACTCTCCACCCCTTCCCTGTGCTGAGGGGCTTCAACACCTCTCTCT 2650
QY 852 AspLeuArgLeuSerProLeuThrProSerThrValArgLeuHisTyrCysProProThr 871

```
Db 2649 GACCTGCGCTGAGCCCTCGACCATCCACCGTTTCGGTTACACTGGTGTCCGCCACG 2590
Qy 872 GluProAsnGlyGluIleValGluTyrLeuIleLeuTyrSerAsnAsnHisThrGlnPro 891
Db 2589 GAGCCCAATGGGAGATGGAGTATCTAAATCTCTACAGCAACACACACCCAGCC 2530
Qy 892 GluHisGlnTrpThrLeuLeuThrGlu----- 901
Db 2529 GAACACCACTGGACACTGCTCACCACAGA -GGTAAGGAGCGCCACCCCTCGCTCCTTGC 2471
Qy 901 ----- 901
Db 2470 CCTTTGTTCCAAGGCTCTCGGATTTGGAGGCTGAAGCTACTGTCTGGGTGCAAGATTG 2411
Qy 902 -----Gly 902
Db 2410 GTTTGCCCATGGTGCTTTGGCTCTCGCCACCCCTGTCTCTCCCACTTCCCAACACAGA 2351
Qy 903 AsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePheIysMet 922
Db 2350 AACATCTTCAGTGCAGAGTCCATGGCTTAGAGAGTGACACTCGGTATTTCTTCAAGATG 2291
Qy 923 GlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIleThrLeu 942
Db 2290 GGAAGCCCGCACAGAGTGGGGCCCTTTCCCGCTTTCAGGATGTGATTACTCTG 2231
Qy 943 GlnGluThrPheSer----- 947
Db 2230 CAAAGACATTTCTCAGGTACCAGGAGGAGGAGAGGCCCTTTGGGACAACTGGAGGCC 2171
Qy 947 ----- 947
Db 2170 CAAGACCCCTGGATTGGCTCCCAAGGCCCTGCTCTCTGGAAATTACATTTTCTTTACTTT 2111
Qy 947 ----- 947
Db 2110 GGAACCTTAAGGTGTGGGGATGGAGAAGCAGTGACAGCTTGATAGAGGAGCTGTCTATAG 2051
Qy 947 ----- 947
Db 2050 CTTTGAGGGGTCTTTTATTATTTTATTCTCTCTGTGTGTGTGTGGCAGATGATATGCG 1991
Qy 947 ----- 947
Db 1990 AAGGTCAGAGGACATCAGAGATTGGTTTGTGGGGATCTTGTCTTGGCTTGGAGGTTGG 1931
Qy 947 ----- 947
Db 1930 GTGAGCTCCTGAGTTCTGTGAAGAAGTGGTCAATAGGCTGGATGAGACACTCAAGGGCT 1871
Qy 947 ----- 947
Db 1870 CAATTAAGATACAAAATAGATCCAGGGCTCCTTATTGTAAATTATTGAAGATGTGAAG 1811
Qy 947 ----- 947
Db 1810 ATGAGCCAGTGGGGCATTAGCCACACATAGGACTGAGGAGCATGAAGCCCTGAGTCCAAA 1751
Qy 947 ----- 947
Db 1750 TGCACACACACACACACACACACACACACACACACACCGTAAAACTTGGCTGTGGC 1691
Qy 947 ----- 947
Db 1690 TGCTGTAATCCTTGTCTAGGATGCAGACAGGATATTTCTGGGGCTGAGAGGACTGGG 1631
Qy 947 ----- 947
Db 1630 GAAGTCAGTGATGTAGACCTCTGATCTCCACTACATATATACATGCCCTTAGATGACA 1571
Qy 947 ----- 947
Db 1570 TATACAGACATGACACACACATGTGTGCACCTAGAGCCACCTTGGTTATACATATATAC 1511

Qy 947 ----- 947
Db 1510 ACCATGAGAGCCCTTGTGTGCCAGGAGAAACAACCTTCACTGAAGAGCTGATTCTATAGA 1451
Qy 947 ----- 947
Db 1450 GTCCAGTTTGTGAGGAGTGGCAGGTCTGTACAGAGGAGAACCGTGTGTGACTCAGT 1391
Qy 947 ----- 947
Db 1390 ATGAGGCCCAAGTAAGATCCCGTGGCTCTCTCTATTTCCAATCAGAAGGAGTGAGCTTCCA 1331
Qy 947 ----- 947
Db 1330 AACCAAGTGTGAACCAAGGTGTGTACTGCACACCTTTTGATCTCCGTCTCTGTATTATGAGC 1271
Qy 948 -----AspSerLeuAspValHis 953
Db 1270 ACTGCTTCTTCTGACCTTGGAGTCACTGTTCTCCCGCGCAGACTCTCTTGGATGTGCAC 1211
Qy 954 AlaValThrGlyIleIleValGlyValCysLeuGlyLeuLeuCysLeuAlaCysMet 973
Db 1210 GCGGTACAGGGGATCATCGTGGGTGTCTGCTGTGGCTTCTCTGCTCTCTGCTGCTGCATG 1151
Qy 974 CysAlaGlyLeuArgGlnSerSer----- 981
Db 1150 TGTGCTGGCTTACGAGGAGCTCCCAAGGTGAACCGGGGAACTGAGACCACTGAGT 1091
Qy 981 ----- 981
Db 1090 TACCAGAAAGGCTGCAGAGTGGACAAAGAGCGTGTCTCTCTGCTGCCCGGAAGCC 1031
Qy 981 ----- 981
Db 1030 CCAGTATAGTAGGCTATGAGAGCTGGAGTGAAGTTTCAGTGTCTACATGTGCCCCCATGTG 971
Qy 981 ----- 981
Db 970 AGGCCCTGGCGTCTCATCTCAGTAGCACAAACAAAGTGGATTAATTTATTATTAGATTTC 911
Qy 981 ----- 981
Db 910 TGAGATGAAGGATTGAGGTGTGTCTCATGAGAATGTAGATAAACTTAGGGAAGCCCC 851
Qy 981 ----- 981
Db 850 AGGACTCCCCACCTCAGTAAACAGCGGGCTGAGTACTGTGGGCTTCCCTCATGCGCTTAA 791
Qy 981 ----- 981
Db 790 ATGGAAATGGCCCATCTCGCCATACAGGAAGGTGCCATCTGAGGGGTTCAGGCCGAGTG 731
Qy 981 ----- 981
Db 730 GACTTGTAFTCTCTTCTCTTGTGATTGTTGACCAACACAGGAAACCAAAATCCACTCTTG 671
Qy 982 -----HisArgGluAlaLeuProGlyLeuSerSerSer 992
Db 670 TCCGACCCCAATAACTTTGTTCTGTGATCAAGGGAGCCCTTCCCGGATTTGCTCTCTCA 611
Qy 993 GlyThrProGlyAsnProAlaLeuTyrThrArgAlaArgLeuGlyProProSerValPro 1012
Db 610 GGCACCCCAAGAAACCCAGCGCTTACACAAAGAGTTCGACTTGGGCCCCCGAGTGTCCCT 551
Qy 1013 AlaAlaHisGluLeuGluSerLeuValHisProArgProGlnAspTrpSerProProPro 1032
Db 550 GCTGCGCATAGTGTGGAGTCTCTGTGATCTCTGTGCTCTCTGTGCTGCTGCTGCTGCTGCT 491
Qy 1033 SerAspValGluAspIysAlaGluValHisSerLeuMetGlyGlySerValSerAspCys 1052
Db 490 TCAGATGTGGAAGACAAGGCTGAAGTACACAGCCTTATGGGTGGCAGTGTTCAGATTGC 431
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 14, 2003, 06:11:05 ; Search time 3845 Seconds
(without alignments)
7913.967 Million cell updates/sec

Title: US-09-754-997A-2
Perfect score: 6604
Sequence: 1 MARADRGRLGLVLTFCLLSA.....CPLTVSPSLPAPVSSAQVP 1252

Scoring table:
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Rgapop 6.0, Rgapext 7.0
Delop 6.0, Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09754997/runat_14102003_061100_9718/app.query.fasta_1.1415
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09754997 @CGN 1 1 2865 @runat_14102003_061100_9718 -NCPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2595	39.3	2626	11	AK079073	Mus muscu
2	1610.5	24.4	4185	11	AK083540	AK083540 Mus muscu
3	1326	20.1	900	12	B1185886	UNL-P-FN-
4	1292	19.6	817	14	CA317106	UT-M-FW0-
5	1252	19.0	740	13	BU701299	UT-M-EX0-
6	1216	18.4	699	14	CD351162	UT-M-FY0-
7	1204	18.2	713	14	CB524810	UT-M-PY0-
8	1202	18.2	741	14	CD352400	UT-M-GI0-
9	1174	17.8	719	14	CA751384	UT-M-F00-
10	1098	16.6	631	14	CB518504	UT-M-GH0-
11	1097	16.6	759	14	CB988160	AGENCOURT
12	1082.5	16.4	703	14	BY712424	BY712424
13	1047	15.9	591	13	BU057120	UT-M-F00-
14	1046	15.8	2906	11	AK051027	Mus muscu
15	1032	15.6	651	10	BB624536	BB624536
16	1008	15.3	637	10	BB665894	BB665894
17	1002	15.2	630	10	BB626002	BB626002
18	1000.5	15.1	3230	11	AK011916	Mus muscu
19	990.5	15.0	717	13	BX100607	BX100607
20	977	14.8	553	14	CA316452	CA316452
21	971	14.7	632	10	BB649997	BB649997
22	960	14.5	544	14	CA892088	CA892088
23	945.5	14.3	1137	11	AK082162	AK082162
24	927	14.0	578	14	CB607425	CB607425
25	910	13.8	595	12	BM256356	BM256356
26	882	13.4	504	14	CA876082	CA876082
27	879	13.3	509	14	CA873325	CA873325
28	803	12.2	455	14	CA883713	B0105A08-
29	784	11.9	442	14	CA891757	CA891757
30	778	11.8	435	14	CA892100	B0170B06-
31	753	11.4	672	9	AW143015	AW143015
32	725	11.0	747	13	B0382531	B0382531
33	708	10.7	438	14	CB751460	CB751460
34	700.5	10.6	836	14	CB994721	CB994721
35	689	10.4	524	9	AA947283	AA947283
36	687	10.4	564	28	AZ354558	AZ354558
37	682	10.3	466	9	AI386698	AI386698
38	678	10.3	400	14	CB699715	CB699715
39	671.5	10.2	411	14	CA871247	CA871247
40	667.5	10.1	3642	11	AK084609	AK084609
41	638.5	9.7	485	9	AI741225	AI741225
42	601	9.1	353	14	CA886377	CA886377
43	588	8.9	624	13	BU056331	BU056331
44	575.5	8.7	500	12	B1182798	UNL-P-FN-
45	564	8.5	939	13	BQ646105	AGENCOURT

ALIGNMENTS

RESULT 1

AK079073

LOCUS

DEFINITION

AK079073

VERSION

KEYWORDS

SOURCE

ORGANISM

AK079073

GI:26098261

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AK079073 2626 bp mRNA linear HTC 05-DEC-2002
Mus musculus adult male diencephalon cDNA, RIKEN full-length
enriched library, clone:933015614 product:hypothetical protein,
full insert sequence.


```

Db      495  TTTGAAATGCCACACCAAGGGGCTTCCAGCCCCCATCATTAATTGGGAAAGGACCAAGGTG 554
Qy      181  ThrValProGluGluProArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuAsp 200
Db      555  ACCGTGCTGAGGAGTCCCGGCTCATCATCTTCCCAATGGCGTCTCCAGATCTTAGAT 614
Qy      201  ValGlnAspSerAspAlaGlySerTyArgCysValAlaThrAsnSerAlaArgGlnArg 220
Db      615  GTCCAGGACAGTATGATGAGGCTCTCACTCGGTGCGTCCACCAATTGAGCCGCCAACGA 674
Qy      221  PheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGly 240
Db      675  TTCAGCAGAGGGCTCGCTCACTGTCGGCCCTTTAAAGGGTCTTTGGGGGGTAAACAAGGA 734
Qy      241  GlnAspValValIleVal-AlaAlaProGluAsnThrThrValValSerGlyGlnAsnVa 260
Db      735  CAGGATGTGTCATTTGTGCCAGCCCGCAGAGAACCAACCGTAGTGTCTGACAGAGTGT 794
Qy      260  lValMetGluCysValAlaSerAlaAspProThrProPheValSerTrpValArgGlnAs 280
Db      795  AGTGATGAGTGGCGTGGCTCTGCTGACCCACCCCTTTTGTGTCTGGGTCCGACAGGA 854
Qy      280  PGLyIleProIleSerThrAspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSe 300
Db      855  TGGAAAGCCCTATCTCCAGGATGTGTCATGTTCTGGGCGGACCAATCTACTCTCCAG 914
Qy      300  rAlaGlnProArgHisSerGlyVal-TyrValCysArgAlaAsnLysProLeuThrArgA 320
Db      915  CGCCAGCCCTCGGACCTTTGAAATTTATTTTCCCAACCCACCAACCCCTTTCCAGCCTT 974
Qy      320  spPheAlaThrAlaAlaGluLeuArgValLeuAlaAlaProAlaIle-SerGlnAla 339
Db      975  AATTTCCTCCCTTCGGCTTCTTAAGTTCCAATTTCTTCTCCCCACCCCTTTTCCAGCCC 1034
Qy      340  ProGluAlaLeu-SerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGl 359
Db      1035  GCCCAAGCCCTTTTCCGAAACCCCGGCCACCCCGCCCTTCTTTTCCCGGCTTCCGG 1094
Qy      359  YGluProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgProAsnGlyAr 379
Db      1095  GGACCCCGCGCGCTGCACTGGCTGCAGCAGCGAATCCGTTGGACCCCAATGGGCG 1154
Qy      379  gValLysValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAl 399
Db      1155  CGTCAAGGTGCAGGCGGTGGCGGACGCTGGTCTCATCTCAGATCGGCTTCAGGACGC 1214
Qy      399  aGlyTyTyThrGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaProLe 419
Db      1215  TGGCTACTACGAGTGCCTAGCAGAAACAGCGCGGAACCTGCCTGTGCGCTGCGCCCT 1274
Qy      419  uAlaValValValArgGluGlyLeuProSerAlaProThrArgValThrAlaThrProLe 439
Db      1275  GCGGATAGTGTGGCGAGGGGCTGCGCGCGCCCGCACTCGGTCCAGCCACGCGCT 1334
Qy      439  userSerSerValLeuValAlaTrpGluArgProGlu-LeuHisSerGluGlnIleI 459
Db      1335  GAGCAGTCTCTGTGTGTGGCTGGAGCGGCTGAGTTTGACAGGAGCAATCA 1394
Qy      459  leGly-PheSerLeu-HisTyThrGlnLysAlaArgGlyValAspAsn-Val-GluTyThrGln 477
Db      1395  TTGGCTTCTCTCTTCTTCTTACCAAAAGGCAAGGGAGTGACAAATGTTGGAGTACAG 1454
Qy      478  PheAlaVal---AsnAsnAspThrThrGluLeuGlnValArgAspLeuGlu----- 493
Db      1455  TTTGCACTAAACCAAAATGGGAAACCAAA-----AGAGCTTGGAGGGTTCGG 1505
Qy      494  -----ProAsnThrAsp---TyrGluPheTyThrValValAlaTyr---Ser 506
Db      1506  GGAACCTGGGAAACCAACCCCGGATTAATGAAGTTTAAAGTGTGGCTTAAGTCTCC 1565
Qy      507  GlnLeuGlyAla---SerArg-ThrSerSerPro---AlaLeuValHisThrLeuAspAs 524
      ::::|||||

```

```

Db      1566  AAGCTGGGGGGCCAGGGCGGAACCTTCAAGCCCGCCCTGGTGATACACTGGACGA 1625
Qy      524  pValProSerAlaAlaProGlnLeuThrLeuSerProAsnProSerAspIle-Argv 544
Db      1626  TGTCTCCCGACGCGACACCCAGCTTACCTTGTTCAGGCCCAACCCCTCGACATCCAGG 1685
Qy      544  alAlaTrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLysTyTyIleG 564
Db      1686  TGGCATGGCTGCCCTTCCCATGATGAGCAATGACAGAGTGTCTGAAAGTCAAGATAG 1745
Qy      564  LuTyRGLyLeuGlyLys---GluAspGlnValPheSerThrGluValProGlyAsnGluT 583
Db      1746  AGTACCGGTTTGGGAAAGGAAGATGAGTTTCTCCACCGAGGTGCTTGGAAATGAAA 1805
Qy      583  hrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyArgValArgIleSerAlaG 603
Db      1806  CACAACCTTACGTTAACTCACTTCAGCCCAACAAAGTGTACCGAGTCCGAGTTTCACTG 1865
Qy      603  lyThrGlyAla-GlyTyRGLyValProSerGlnTrpMetGlnHisArgThrProGlyVal 622
Db      1866  GCATCGCGCTGGGCTATGAGTCCCTTCTCAGTGTGATGCAGCAGACGACACCTGGTGTG 1925
Qy      623  HisAsnGlnSerHisValProPheAlaProAlaGluLeuLys----- 636
Db      1926  GACAACACAGGACCATGTT-----ATGGAACATAAAGGGGCGAGAATCGAAAGGG 1973
Qy      637  -----ValArgAlaLysMetGluSerLeuValValSer-----TrpGln 649
Db      1974  GGGACAGGTGCTCTGTGGGCTTAGATATTGGCCCTTTTGGCTGTTATTGGCAA 2033
Qy      650  ProProPro 652
Db      2034  CCCCCCCCC 2042

RESULT 2
AK083540
LOCUS
DEFINITION
Mus musculus 9 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:D030042G07 product:hypothetical Fibronectin
type III domain containing protein, full insert sequence.
ACCESSION
AK083540.1 GI:26350620
VERSION
HTC; CAP trapper.
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
PUBMED
10349636
REFERENCE
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
PUBMED
11042159
REFERENCE
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K.; Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
JOURNAL
MEDLINE

```

PURMED
JOURNAL
AUTHORS

11076861
4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Aizawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Ozaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Aeburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nakaide, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Stauble, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. P., Brownstein, M. J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzei, J., Mombert, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. P., Suzuki, H.,
Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.,
and Hayashizaki, Y.

TITLE
JOURNAL
MEDLINE
PURMED
REFERENCE

Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851

PURMED
JOURNAL
AUTHORS

5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 4185)

JOURNAL
REFERENCE
AUTHORS

Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Aizawa, K., Hanagaki, T., Hara, A., Hashizume, M.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tonaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M., and Hayashizaki, Y.

TITLE
JOURNAL

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC).
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

FEATURES
source

Location/Qualifiers
1. .4185
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:D030042G07"
/clone="D030042G07"
/clone_xref="txon:10090"
/tissue_type="whole body"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="9 days embryo"
136. .3195
/note="unnamed protein product; hypothetical Fibronectin

CDS

type III domain containing protein (InterPro|IPR003961,
evidence: InterPro)
putative"
/codon_start=1
/protein_id="BAC38947.1"
/db_xref="GI:26350621"
/translation="MDRVATLPSGVLIQYDVDPDAGNYRCVAATIAHKRKSMEASLT
IVANERSFYMTIIASPNVNTASLHQTQVLECMATGYPRPIISMRLOHKSIDVEN
TRVLGNGLIISVKLOHAGYVCRATPTGTRNTVAMAILTVLAPSPFVEMPSLFR
PRAGTARFVCOAGIPSPKMSLWKNRIHNSGRIKYNLSKLVINQIIPEDDAILQCM
AENSQGVLSARLTIVMSDRPSAPNTVHAETMSSAILLAWERPLYSKSDIYASV
HYMKAGLNNEEYQVNLGNDTHYIIDLEPDSNTFYIVAYMPLGASQMSDRVTQNT
LEDVLPPEISLTSRPTDLVSMPLIPAKYRGQVLYRLSVKSPKASPELHLEP
GTWHEYLGLKDPDVLVRIATATRVGLGESSVMTSHRTPKATSKVAPKSPLEHLEP
LNCITISVRHLQDTEPAARIGYKLPYKRGQQRHGPIFLDTGDLITLSGLDPRKY
HVRLLAVNMEEGYQAOVTSTPCVSRVRMVPVPPPHHLYAKANTISFIFLHWR
PAFTTAQYNTYICNVGLQNSALVLYLQTSETHLVQGLEPNTKYEFVRLHVDQL
SSPSPVYHSTLPEAPTGPVGVKVLIEDDTALVSKPKPDGPTVVTVRTYILYASR
KWTAGMQLVLRREGAITMLLENLVAGNVYIKI SASNEVGEPPNSVELAVLPKD
ASESNQPKLSDSNKAVSYGYHLDOKSMVTGIAGVGIALTCILICVLILVSKAR
KSSASKTAQSGTQPLQSASASVAAGSMGNKLERATETASLSVPMMSFIDAKGGTD
LIINSYGIINKNNTKKKLPFDTKIKVEQTRRFTQAVCFYQGTIVTLISDSDSPG
SPQGTASFPFPFGATLDTBHSANSSEHSGTSDSRFSHESNDIHLSSVISSTPTPS
NPLAGSDSDGAAPKKGDPAPLPAPETQTSAPQTSAGLRVAAGFPV"

BASE COUNT 1088 a 1049 c 1029 g 1019 t
ORIGIN

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	4185
3.26e-98	1610.50	Matched:	357	
Percent Similarity:	55.37%	Conservative:	153	
Best Local Similarity:	38.76%	Mismatches:	344	
Query Match:	24.39%	Gaps:	67	
DB:	11	Indels:	15	

US-09-754-997a-2 (1-1252) x AK083540 (1-4185)

QY	144	PheSerLeuHisProGluSerGlnIleValGluGluLeuGlyThrAlaArgPheGluCys	163
DB	10	TTTGAAGTCCATCCAGTGTCTACTGAGGTCTCCATGAGGAGGAGTGTGTAGATTTTCATGCG	69
QY	164	HisThrLysGlyLeuProAlaProIleThrTrpGluLysAspGlnValThrValPro	183
DB	70	AAGATTTCATCAACGCCCTGCGAGTCTATACATGAGGAGTATTAATAGAACACCCCTGGCT	129
QY	184	GluGlu---ProArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuAspValGln	202
DB	130	ACAACCATGGACGGGTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	189
QY	203	AspSerAlaAlaGlySerTrpArgCysValAlaThrAsnSerAlaArgGlnArgPheSer	222
DB	190	CCGAGGAGTGTGGGATTTACCGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	249
QY	223	GlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAsp	242
DB	250	ATGGAGGCTTCGCTAACCATTA-----GTTCAGCTAATGAGAACAGATCC	294
QY	243	Val-----ValIleValAlaAlaProGluLeuThrValValSerGlyGlnAsn	259
DB	295	TTCTACATGCCAACCATTTATAGCCAGTCCACAGATGTAACCGCATCTCTACATCAGACA	354
QY	260	ValValMetGluCysValAlaSerAlaAspProThrProPheValSerTrpValArgGln	279
DB	355	GTGTGTGTAGAGTGTATGGCCACAGGATATCCAGACCCCATCATTTCTCGAGCGCTCTT	414
QY	280	AspGlyLysProIleSerThr---AspValIleValLeuGlyArgThrAsnLeuLeuIle	298
DB	415	GATCAAAAGTCCATGTGTCTTCAACACCTCGGGTACTTGGAAATGGCAACCTCATATA	474
QY	299	AlaSerAlaGlnProArgHisSerGlyValTrpValCysArgAlaAsnLysProLeuThr	318
DB	475	TCAGATGTCAGCTGCAGCATGCTGGAGTGTATGTTGTGGGCCACCCACCCCGGCACA	534

Qy 319 ArgAspPheAlaThrAlaAlaGluLeuArgValLeuAlaAlaProAlaIleSerGln 338
|||
Db 535 CGTAACCTTACAGTTGCTATGCGCAACGTTAACTGTTATAGCTCTCTCTTCAATTGTCGAA 594
|||
Qy 339 AlaProGluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSer 358
|||
Db 595 TGGCCTGAAAGTTTAAACAGGCCACGAGCTGGCACTGCGACGATTTGTGTGCGCAAGAA 654
|||
Qy 359 GlyGluProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgProAsnGly 378
|||
Db 655 GGAATTCCTCACCCAAATGCTCTGTTGAAATGGAAGAGATACATTCAATGGC 714
|||
Qy 379 ArgVallysValGlnGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAsp 398
|||
Db 715 AGGATTAATGTAC-----AACAGCAATTTGGTCATTAACTAGATTATCCCGGAAGAC 768
|||
Qy 399 AlaGlyTyrrGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaPro 418
|||
Db 769 GATGCTATCTATAGTCATGCTGGAACAGCCCAAGGATCTGTCTGTCGAGCCAGA 828
|||
Qy 419 LeuAlaValValAlaArgGluGlyLeuProSerAlaProThrArgValAlaThrPro 438
|||
Db 829 CTGACTGTGGTAATGTCGGAAGATAGACCCAGCGCTCTTACACGTTTCATGCTGAGACC 888
|||
Qy 439 LeuSerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGlnIle 458
|||
Db 889 ATGTCGAGCTCTGCTATCTTCTGCTTGGGAGAGACCCTGTACAATTTCAGACAGGTC 948
|||
Qy 459 IleGlyPheSerLeuHisTyrrGlnIleAlaArgGlyValAspAsnValGluTyrrGlnPhe 478
|||
Db 949 ATCCCTACTCTGTACACTATACGAAGCGGAAGGTTTAAATAACGAGGATATACAGGTG 1008
|||
Qy 479 AlaValAsnAsnAspThrGluLeuGlnValArgAspLeuGluProAsnThrAspTyr 498
|||
Db 1009 GTCCTTGAAGACACACACTCATATCATGATGACTTGTAGAACCCGACGCACTAT 1068
|||
Qy 499 GluPheTyrrValAlaTyrrSerGlnLeuGlyAlaSerArgThrSerProAlaLeu 518
|||
Db 1069 ACTTCTACATTTGGCGTACATCGCGCTGGGGCCAGCCAGATGTCGATCATGTGACA 1128
|||
Qy 519 ValHisThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSerProAsn 538
|||
Db 1129 CAGAACACGCTAGAGATGTGCCCTTAAGACCTCCCGAAATTAGCTTGACGATCGAAGT 1188
|||
Qy 539 ProSerAspIleArgValAlaTrpLeuProLeuProSerSerLeuSerAsnGlyGlnVal 558
|||
Db 1189 CCCACCGACATCTCTGCTCTGCTGCCAATCCCGGCCAAGTACCAGCGGGGCCAAGTG 1248
|||
Qy 559 LeuLysTyrrLysIleGluTyrrGlyLeuGlyLysGluAspGlnValPheSerThrGluVal 578
|||
Db 1249 GTGCTGTACCGCTGTCTTCCGCTTGAGCACTGGAAGCCATCCAAAGTTGTAGAACTC 1308
|||
Qy 579 ProGlyAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyrrArgVal 598
|||
Db 1309 CCGGGGACCGTCAGCAGTACCTTCTGGAAGCCCTGAAGCCAGACAGTGTCTATCTGGTT 1368
|||
Qy 599 ArgIleSerAlaGlyThrGlyAlaGlyTyrrGlyValProSerGlnTrpMetGlnHisArg 618
|||
Db 1369 CGGATTACTGACGCTACCCGTGGTGGCTGGGAGAGTCTCGTGGTGGACCTTCACACAGG 1428
|||
Qy 619 ThrProGlyValHisAsnGlnSerHisValProPheAlaProAlaGluLeuLysValArg 638
|||
Db 1429 ACGCCCAAGGCT---ACGAGCGTGAAGACCCCGAAGTCTCT---GAACCTGCATTTGGAA 1482
|||
Qy 639 ---AlaLysMetGluSerLeuValSerTrp---GlnProProHisProThrGln 656
|||
Db 1483 CCACCTGAATTTGACTACCTCTCTGTGAGGTGGCTGCAAGATACAGAGACCCCGCCGCT 1542
|||
Qy 657 IleSerGlyTyrrLysLeuTyrrTrpGlyGluValGlyThrGluGluAlaAspGlyAsp 676
|||
Db 1543 ATTCGGGGCTACAAGCTGTTTATTAAGAGGAGGCGCAGCAAGAACAC----- 1590
|||
Qy 677 ArgProProGlyGlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLysLys 696
|||

Db 1591 -----|||
|||
Qy 697 LysVallysGlnTyrrGluLeuThrGlnLeuValProGlyArgProTyrrGluValLysLeu 716
|||
Db 1612 GGTGACCTGCTGTATAGCTCAGTGGCTTGAGCCCGAAGAAATAACACGAGGAGCTC 1671
|||
Qy 717 ValAlaPheAsnLysHisGluAspGlyTyrrAlaAlaValTrpLysGlyLysThrGluLys 736
|||
Db 1672 CTGGCTTACAAACAACATGGAAGAGGCTACAGGCC-----GACCAGACT 1716
|||
Qy 737 AlaProThrProAspLeuProIleGlnArgGly-----ProProLeuProPro 752
|||
Db 1717 GTCAGCCTCCCGATGCGTGTGCTGCGTGCATGCGTGTCTCCACCCAGCCGACCC 1776
|||
Qy 753 AlaHisValHisAlaGluSerAsnSerThrSerIleTrpLeuArgTrpLysLysPro 772
|||
Db 1777 CACCATCTCTACGCGAAGGGAATACCTCTCTTCATCTCTCTGCACTGAGAGAGGCT 1836
|||
Qy 773 AspPheThrThrValLysIleValAsnTyrrValArgPheGlyProTrpGlyLeuArg 792
|||
Db 1837 CGGTTACCAACCCCGCAAGTCATTAACTACACCATCGCTGTAACCTGTGGCTGCAG 1896
|||
Qy 793 AsnAlaSerLeuValThrTyrrThrSerSerGlyGluAspIleLeuIleGlyLeu 812
|||
Db 1897 AATGCTCTCTGCTTCTTACTCTGCAACATCAGAAACCCACATGCTGTGTCGAAGGACTA 1956
|||
Qy 813 LysProPheThrLysTyrrGluPheAlaValGlnSerHisGlyValAspMetAspGlyPro 832
|||
Db 1957 GAACCAACACCAAGTATGAATTTGCTGTCGCACTGCATGTGATCATGCTGCCAGTCCC 2016
|||
Qy 833 PheGlySerValValGluArgSerThrLeuProAspArgProSerThrProProSerAsp 852
|||
Db 2017 TGGAGTCTGTGCTTACCATTTCTACATTCGCAAGACCCACAGGCCCCCGGTAGGG 2076
|||
Qy 853 LeuArgLeuSerProLeuThrProSerThrValArgLeuHisTyrrCysProProThrGlu 872
|||
Db 2077 GTAAAGGTGACCTCTAGAGACGACACACTGCCCTGTGTTCTTGAAACCCCAAGATGGC 2136
|||
Qy 873 ProAsnGlyGluIleValGluTyrrLeuIleTyrrSerAsnAsnHisThrGlnProGlu 892
|||
Db 2137 CCAAGACAGTGTGTGACACGCTACCATCTTGTATGCTTCCAGGAAGGCTCGAATGGC 2196
|||
Qy 893 HisGlnTrpThrThrGluGlyAsnIlePheSerAlaGluValHisGlyLeu 912
|||
Db 2197 GGAGGTGGCAGGTTCTTCCCGGAAGGGGCAATCACCATGGCTTGTCTGCGAAACCTG 2256
|||
Qy 913 GluSerAspThrArgTyrrPhePheLysMetGlyAlaArgThrGluValGlyProGlyPro 932
|||
Db 2257 GTGGCAGGGAACGTTACATTTGTCAAGATATCTGCGTCCAACGAGGTGGGAGAGGGCCC 2316
|||
Qy 933 PheSerArgLeuGlnAspValIleThrLeuGlnGluThrPheSerAspSer----- 949
|||
Db 2317 TTTTCAAACTCCGTTGGAGCTGGCGCTCTCCCAAGGATGCTTCAGAGTTCGAACACAGG 2376
|||
Qy 950 -----LeuAspVal 952
|||
Db 2377 CCCAAGGTTGGATTTCTCAATGCCAAAGTTTATTCGGCTATTATACCATCTGACCAA 2436
|||
Qy 953 HisAlaValThrGlyIleIleValGlyValCysLeuGlyLeuLeuCysLeuAlaCys 972
|||
Db 2437 AAGTCAATGAGGCGCATCGCAGTAGGCTTGGCATAGCTTGCCTGTATCTCTCATCTGT 2496
|||
Qy 973 MetCysAlaGlyLeu---ArgGlnSerSerHisArgGluAlaLeuProGlyLeuSerSer 991
|||
Db 2497 GTTCTCATCTTGTATCTACCGAAGCAAGCCAGGAAGTCATCCGCTCTAAGACAGCGCAG 2556
|||
Qy 992 SerGlyThr-----ProGlyAsnProAlaLeuTyrrThrArgAlaArgLeuGly 1007
|||
Db 2557 AGTGAACCCAGCGGTTATCCCAAGCCAGGCTGCTCTGTAGCAGCAGGAGTGATCGGG 2616
|||
Qy 1008 ProProSerValProAlaAlaHisGluLeuGluSerLeuValHisProArgProGlnAsp 1027
|||

Db 2617 AAGAACCTGGAGAGACTGACAGAGACTGCTGGTACCCATGATGCCAGCAGC 2676

QY 1028 Trp 1028

Db 2677 TTC 2679

RESULT 3

LOCUS B1185886 900 bp mRNA linear EST 10-JUL-2001

DEFINITION UNL-P-FN-df-h-02-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone

ACCESSION B1185886

VERSION B1185886

KEYWORDS EST.

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

AUTHORS 1 (bases 1 to 900)

TITLE Caetano, A.R., Johnson, R.K. and Pomp, D.

JOURNAL Generation and sequence characterization of a normalized cDNA library from swine ovarian follicles

MEDLINE Mamm. Genome 14 (1), 65-70 (2003)

PUBMED 22419904

COMMENT 12532269

Contact: Pomp, D

Department of Animal Science

University of Nebraska, Lincoln

Lincoln, NE 68583-0908, USA

Tel: 402 472 6416

Fax: 402 472 6362

Email: dpomp@unl.edu

Oligo-dT track not found, Not 1 site shown in beginning of sequence is likely internal to the message.

Seq primer: M13 -29

POLYA=No.

FEATURES

Source

1..900

/organism="Sus scrofa"

/mol_type="mRNA"

/strain="University of Nebraska, Lincoln Swine Selection Lines"

/db_xref="taxon:9823"

/clone="UNL-P-FN-df-h-02-0-UNL"

/dev_stage="ADULT"

/lab_host="DH10B (Life Technologies)"

/clone_lib="UNL-P-FN"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; The UNL-P-FN library is a normalized library representing porcine ovarian follicles, ranging between 2.0 to 10.0 mm in diameter, collected during 7 days of the follicular phase of the pig estrous cycle. This library was derived from the library UNL-P-F2. The tag is a string of 5-6 nucleotides present between the Not 1 site and the oligo-dT track. The library was constructed as described by Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996.

BASE COUNT 149 a 304 c 291 g 150 t 6 others

ORIGIN

Alignment Scores:

Pred. No.: 8-27e-80 Length: 900

Score: 1326.00 Matches: 262

Percent Similarity: 97.12% Conservative: 8

Best Local Similarity: 94.24% Mismatches: 8

Query Match: 20.08% Indels: 1

DB: 12 Gaps: 0

US-09-754-997A-2 (1-1252) x B1185886 (1-900)

QY 247 AAlaAlaProGluAsnThrThrValValSerGlyGlnAsnValValMetGluCysValAla 266

Db 7 GCTGCCCGCAGAACACCACTGCTGCGGCCAGAGCGTGGTGTAGTGTGAGTGTGCGCC 66

QY 267 SerAlaAspProThrProPheValSerTipValArgGlnAspGlyIysProIleSerThr 286

Db 67 TCGGCTGACCCCACTTATCGTGTCTCGGTCCGACAGACGGAGAGCTATTTCATACC 126

QY 287 AspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSerAlaGlnProArgHisSer 306

Db 127 GATGTCTATCGTCTGCGCGCCGACCACTGCTANTACACGAGCGGCANCCCGG-CATCT 185

QY 307 GlyValTyrValCysArgAlaAsnLysProLeuThrArgAspPheAlaThrAlaAla 326

Db 186 GCGGTCTACGTCTGCAGAGCCAAACAGCCCGCTAGCGCGCACTTGGCACCGCCGCGCC 245

QY 327 GluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerArgThr 346

Db 246 GAGCTCCGCGTGTGCGCGCTCTGCGCATCTCGCAGCGCCCGAGGCGCTATCTCGGAGC 305

QY 347 ArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLeuHis 366

Db 306 AGGCGGAGCACCGCGCTTGTGTGCGCGCAACGGGTGAGCGCGCCGCGCTGGGG 365

QY 367 TrpLeuHisAspGlyIleProLeuArgProAsnGlyArgValIysValGlnGlyGly 386

Db 366 TGGCTGCACAAACGGCGCGCTGCGGCCCAACGGCGGTGTCAAGGTGACGGCGCGCC 425

QY 387 GlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrTyrGlnCysValAla 406

Db 426 AGCAGCCTGGTGCATCAGCGAGATCGGCTGCGAGATGCGGCTACTACCACTGGTGGCC 485

QY 407 GluAsnSerAlaGlyThrAlaCysAlaAlaProLeuAlaValValValArgGluGly 426

Db 486 GAGAACACAGCGCGGCGACCGCGTGGCGCGCGCGCTGGCAGTAGTAGTGGCGAGGGT 545

QY 427 LeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerValLeuVal 446

Db 546 CTGCCAGCGCCCGCAGCGGGTCAACCGCCACCGCCCTGAGCAGCTCCACTGTGCTGTG 605

QY 447 AlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTyrGln 466

Db 606 GCCTGGGAGCGCGCGGAGCTGCACAGCGAGAGATCATCGGTTTCTCTCTTCACTACCA 665

QY 467 LysAlaArgGlyValAspAsnValGluTyrGlnPheAlaValAsnAsnAspThrThrGlu 486

Db 666 AAGGACCGGGCATGGACAGTGTGGATACCAAGTTTGGCGGTGAACATGACACCACTGAG 725

QY 487 LeuGlnValArgAspLeuGluProAsnThrAspTyrGluPheTyrValValAlaTyrSer 506

Db 726 CTACAGGTTTCGGGACCTGGAGCCCAACACAGACTATGAGTTCTATGTGGTGGCTACTCA 785

QY 507 GlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThrLeuAspAsp 524

Db 786 CAGCTGGGGGCGAGCGCCACCTCCACCCCGCTGGTGGTCCACACATGGATGAT 839

RESULT 4

LOCUS CA317106 817 bp mRNA linear EST 26-NOV-2002

DEFINITION UI-M-FW0-cbm-d-21-0-UI.r1 NIH_BMAP_FW0 Mus musculus cDNA clone

ACCESSION IMAGE: 6811462 5', mRNA sequence.

VERSION CA317106

KEYWORDS CA317106.1 GI:24535230

SOURCE EST.

ORGANISM Mus musculus (house mouse)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS 1 (bases 1 to 817)

TITLE NIH-MGC http://mgi.nci.nih.gov/

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

FEATURES

source	Location/Qualifiers	Seq primer: pYX-5.
1. 817	/organism="Mus musculus"	
	/mol_type="mRNA"	
	/strain="C57BL/6"	
	/db_xref="taxon:10090"	
	/clone="IMAGE: 6811462"	
	/tissue_type="whole brain"	
	/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"	
	/lab_host="DH10B (T1 phage resistant)"	
	/clone_lib="NIH BMAP FMO"	
	/notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."	
BASE COUNT	180 a 254 c 148 g 14 t 1 others	
ORIGIN		
Alignment Scores:		
Pred. No.:	1.46e-77	Length: 817
Score:	1292.00	Matches: 245
Percent Similarity:	98.39%	Conservative: 0
Best Local Similarity:	98.39%	Mismatches: 4
Query Match:	19.56%	Indels: 1
DB:	14	Gaps: 0
US-09-754-997A-2 (1-1252) x CA317106 (1-817)		
QY	1004 AlaArgLeuGlyProProSerValProAlaAlaHisGluLeuGluSerLeuValHisPro	1023
Db	2 GCTCGACTTGGGCCCCCAGTGTCCTGCTGCCATGAGTGGAGTCCCTCGGTCATCCT	61
QY	1024 ArgProGlnAspTrpSerProProSerAspValGluAspLysAlaGluValHisSer	1043
Db	62 CGTCCCAGGATGGTCCCAACCCCTCAGATGTGGAAGCAAGCTGAGTACACAGC	121
QY	1044 LeuMetGlyGlySerValSerAspCysArgGlyHisSerLysArgLysIleSerTrpAla	1063
Db	122 CTTATGGTGGCAGTGTTCAGATGGCGGGCCACTCCCAAGAGAAAGATCTCTGGCT	181
QY	1064 GlnAlaGlyGlyProAspTrpAlaGlySerTrpAlaGlyCysGluLeuProGlnGlySer	1083
Db	182 CAGGCGAGGGGACCAACCTGGCAGGCTCTGGGCGAGGCTGAGTGGCCCCAGGGTAGT	241
QY	1084 GlyProArgProAlaLeuThrArgAlaLeuLeuProProAlaGlyThrGlyGlnThrLeu	1103
Db	242 GGTCCAGGCGGCTGACCCCTGCTCTGCTCCCTCCAGCGGAGACCGGGCAGACTG	301
QY	1104 LeuLeuGlnAlaLeuValTrpAspGlyIleLysSerAsnGlyArgLysLysProSerPro	1123
Db	302 CTGCTGCAAGCCCTGGTATATGACGCATTAAGAGCAACGGGAGAAAGAGCGCTCCCA	361

1124 AlaCysArgAsnGlnValGluAlaGluValIleValHisSerAspPheGlyAlaSerLys 1143
362 GCCTGCGAGGAATCAGGTGGAAGCTGAGGTCAATGTCCACTCCGACTTCGTCGATCCAAA 421
1144 GlyCysProAspLeuHisLeuGlnAspLeuGluProGluGluProLeuThrAlaGluThr 1163
422 GGATGTCTTGAGCTCCACCTCCCAAGACCTGGAGCCAGAGAACCACTGATGTCGAGAGACT 481
1164 LeuProSerThrSerGlyAlaValaAspLeuSerGlnGlnValaAspTrpLeuGlyArgGlu 1183
482 CTGCTTCCACGCTCTGGAGCTGTGGATCTGTCTCAAGGAGCAGAGCTGGTGGCAGGGAG 541
1184 LeuGlyGlyCysGlnProThrThrSerGlyProGluArgLeuThrCysLeuProGluAla 1203
542 CTGGGAGGGTGGCAACCAACCAACCAAGCTGGGCGCAGAGAGCTCACCTGCTTGCAGAGCA 601
1204 AlaSerAlaSerCysSerCysSerAspLeuGlnProSerThrAlaIleGluAlaPro 1223
602 GCGAGTGCTCTCTGCTCTGCTCAGACCTCCAGCCCGAGCAGCTATAGAGGAGGCCCT 661
1224 GlyLysSerCysGlnProLysAlaLeuCysProLeuThrValSerProSerLeuProArg 1243
662 GGGAAAGCTGCAGGCCCAAGCCCTGTGTCTCTTAACAGTCAGGCCCAAGCTTCCAGN- 720
1244 AlaProValSerSerAlaGlnValPro 1252
721 GCCCTGCTCTCTGCTCAAGTCCC 747

RESULT 5

BU701299 740 bp mRNA linear EST 09-OCT-2002
UI-N-EXO-bwy-1-18-0-UI.r1 NIH BMAP_EXO Mus musculus cDNA clone
IMAGE: 5704793 5', mRNA sequence.
BU701299
BU701299.1 GI:23624921
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 740)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

FEATURES

source	Location/Qualifiers
1. 740	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="C57BL/6"
	/db_xref="taxon:10090"
	/clone="IMAGE: 5704793"
	/tissue_type="whole brain"
	/dev_stage="embryo 15.5 dpc"
	/lab_host="DH10B (T1 phage resistant)"
	/clone_lib="NIH BMAP EXO"
	/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded

Db 542 GTCCCTGCTGCCCATGAGTTGGAGTCCCTCGTGCATCTCTGTCNCAGAGTTGGTCCCA 601

QY 1031 ProPserAspValGluAspLysAlaGluValHisSerLeuMetGlyGlySerValSer 1050
 |||||
 Db 602 CCACCTCAGATGTGAAGACAGAGGCTGAAGTACACAGCCTTATGGTGGCAGTTTCA 661
 |||||

QY 1051 AspCysArgGlyHisSerLysArgLysIleSerTrpAlaGlnAlaGlyGly 1067
 |||||
 Db 662 GATTGCGGGGCCACTCCANGAGAAAGATCTCTGNGCTCANGCAGGGGA 712
 |||||

RESULT 8
 CD352400 741 bp mRNA linear EST 29-MAY-2003
 LOCUS UI-M-GIO-cgj-p-07-0-UI.r1 NIH BMAP_GIO Mus musculus cDNA clone
 DEFINITION IMAGE: 6856280 5', mRNA sequence.

ACCESSION CD352400
 VERSION CD352400.1 GI:31143911
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 741)

REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefi.html>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.
 Location/Qualifiers
 1..741
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE: 6856280"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dr
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is AGGAGACAG. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

BASE COUNT 139 a 238 c 219 g 145 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1,566-71 Length: 741
 Score: 1202.00 Matches: 240
 Percent Similarity: 98.37% Conservative: 2
 Best Local Similarity: 97.56% Mismatches: 4
 Query Match: 18.20% Indels: 0
 DB: 14 Gaps: 0

US-09-754-997A-2 (1-1252) x CD352400 (1-741)

QY 106 GluAlaLeuArgIleTrpLysValThrGluGlySerTyrSerCysLeuAlaHisSerPro 125
 |||||
 Db 2 GAAGCTCTTAGGATCTGAAAGGTCACTGAGGGCAGCTATTCTCTGTCGCGCCACAGCCG 61
 |||||

QY 126 LeuGlyValValAlaSerGlnValAlaValValLysLeuAlaThrLeuGluAspPheSer 145
 |||||
 Db 62 CTAGGAGTGGTGGCCAGCAGGTTGCTGTGTCAAGCTTGCCACACTCGAAGACTTCTCT 121
 |||||

QY 146 LeuHisProGluSerGlnIleValGluGluLysGlyThrAlaArgPheGluCysHisThr 165
 |||||
 Db 122 CTGACCCCGAGTCCAGATTGTGAGAGAGACCGGACAGCACGCTTTGAATGCCACACC 181
 |||||

QY 166 LysGlyLeuProAlaProIleIleThrTrpGluLysAspGlnValThrValProGluGlu 185
 |||||
 Db 182 AAGGCGCTTCAGCGCCCATCATTTACTTGGGAAAGACACAGGTGACCGTCCCTCAGAG 241
 |||||

QY 186 ProArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuAspValGlnAspSerAsp 205
 |||||
 Db 242 TCCCGGCTCATCACTCTTCCCAATGGCGTCTCTCCAGATCTTAGATGTCCAGGACAGTGAT 301
 |||||

QY 206 AlaGlySerTyrArgCysValAlaThrAsnSerAlaArgGlnArgPheSerGlnGluAla 225
 |||||
 Db 302 GCAGGCTCTACCGCTGCGTGGCCACCAATTCAGCCGCCCAACGATTCAGCCAGGAGGCC 361
 |||||

QY 226 SerLeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAspValValIle 245
 |||||
 Db 362 TCGCTCACTGTGGCCCTCAGAGGCTCTTGGAGGCTACAGGGGGGAGGATGTGGTCATT 421
 |||||

QY 246 ValAlaAlaProGluAsnThrThrValValSerGlyGlnAsnValValMetGluCysVal 265
 |||||
 Db 422 GTGGCAGCCCCAGAGNACACCACTGAGTGTCTGGACAGAGTGTAGTGTGAGTGCCTG 481
 |||||

QY 266 AlaSerAlaAspProThrProPheValSerTrpValArgGlnAspGlyLysProIleSer 285
 |||||
 Db 482 GCCTCTGCTGACCCACCCCTTTTGTGTCTGGGTCCGACAGGATGGAAACCTATCTCC 541
 |||||

QY 286 ThrAspValIleValLeuGlyValArgThrAsnLeuLeuIleAlaSerAlaGlnProArgHis 305
 |||||
 Db 542 ACGGATGTCATCGTTCTGGCGCGGACCAACATCTACTTCATCGCCAGCGCCAGCCTCGGC 601
 |||||

QY 306 SerGlyValTyrValCysArgAlaAsnLysProLeuThrArgAspPheAlaThrAlaAla 325
 |||||
 Db 602 TCTGAGTCTATGTCGCGGAGCAACCAAGCCCCCGACGCGTGTATTTCGCCCACTGCGCT 661
 |||||

QY 326 AlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerArg 345
 |||||
 Db 662 GCTGAGCTCCGAGTGTCTGCTGCCCGCCAGCCATCTCGCAGGCGCCCGAGCGCTCTCGCGG 721
 |||||

QY 346 ThrArgAlaSerThrAla 351
 |||||
 Db 722 ACGCGGGCCAGCACCGCG 739
 |||||

RESULT 9
 CA751384 719 bp mRNA linear EST 27-NOV-2002
 LOCUS UI-M-F00-cdn-j-21-0-UI.r1 NIH BMAP_F00 Mus musculus cDNA clone
 DEFINITION IMAGE: 6830806 5', mRNA sequence.

ACCESSION CA751384
 VERSION CA751384.1 GI:25579257
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 719)

REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

FEATURES

Location/Qualifiers

1..719

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE: 6830806"

/tissue_type="whole brain"

/dev_stage="embryo 12.5dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH BMAP_F00"

/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;

Site 2: Not I; The library was constructed according

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

with EcoR I adaptor, digested with NotI and then cloned

directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

is TGAGAGACC. This library was created for the University

Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the

Developing Mouse Nervous System', supported by National

Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,

Program coordinator."

program coordinator."

BASE COUNT 112 a 234 c 235 g 135 t 3 others

ORIGIN

Alignment Scores:

Pred. No.: 1..18e-69 Length: 719
Score: 1174.00 Matches: 234
Percent Similarity: 98.33% Conservative: 1
Best Local Similarity: 97.91% Mismatches: 4
Query Match: 17.78% Indels: 0
DB: 14 Gaps: 0

US-09-754-997a-2 (1-1252) x CA751384 (1-719)

Qy 225 AlaSerLeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAspValVal 244
Db 1 GCCTCGCTCACTGTGGCCCTCAGAGGCTTTTGGAGGCTACACGGGGGCGAGATGTGTC 60
Qy 245 IleValAlaAlaProGluAsnThrThrValValSerGlyGlnAsnValValMetGluCys 264
Db 61 ATTGTGGCAGCCCGCAGAGACACACCGGTAGTGTCTGGACAGAGTGTAGTATGAGGTGC 120
Qy 265 ValAlaSerAlaAspProThrProPheValSerThrValArgGlnAspGlyLysProile 284
Db 121 GTGGCCTCTGTGACCCCAACCCCTTTGTGTCTGGTCCGACAGGATGGAAACCTATC 180
Qy 285 SerThrAspValIleValLeuGlyValArgThrAsnLeuLeuIleAlaSerAlaGlnProArg 304
Db 181 TCCAGGATGTCTGCTTCTGGCGGACCAATCTACTCTATCGCCAGCCGCGAGCTCGG 240
Qy 305 HisSerGlyValThrValCysArgAlaAsnLysProLeuThrArgAspPheAlaThrAla 324
Db 241 CACTCTGGAGTCTATGTCTGGCAGCACAACAGCCCGCCGCGGTGATTTCGCCACTGCG 300
Qy 325 AlaAlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeuSer 344
Db 301 GCTGTGAGCTCCGAGTGTCTGTCGCCAGCCATCTCGCAGGCGCCGAGGCGCTCTCG 360

Qy 345 ArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgProAla 364
Db 361 CGGACGGCGGCCAGACACCGCGCGCTTGTGTGCGGCGGTCCGGGGAGCCACGCGCCGCG 420
Qy 365 LeuHisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgValValValGlnGly 384
Db 421 CTGCACCTGGCTGATGACGGGATCCCGTTCGACCCCAATGGGCGCGTCAAGGTGCGGGC 480
Qy 385 GlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrTyrGlnCys 404
Db 481 GGTGGCGCGCAGCTTGGTCACTCAGATCGGCGCTTCAGACGCTGGCTACTACCACTGC 540
Qy 405 ValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaAlaProLeuAlaValValValArg 424
Db 541 GTAGCAGAAACACAGCGCGGAACTGCTGTGCGCGCTGCGCCCTCGGCGGTANTGTGTCGC 600
Qy 425 GluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerVal 444
Db 601 GAGGGGTGCGCCANCGCCCGACTCGGTCACAGCAGCGCGCTGACAGCTCTCTCTGTG 660
Qy 445 LeuValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeu 463
Db 661 CTGCTGCGCTGCGAGCGGCGCTGAGTTGACAGCGGCAATCATTTGCTCTCTCT 717

RESULT 10
CB518504 631 bp mRNA linear EST 28-MAR-2003
LOCUS UI-M-GHO-cec-h-04-0-UI.r1 NIH BMAP_GHO Mus musculus cDNA clone
DEFINITION IMAGE: 6837653 5', mRNA sequence.
ACCESSION CB518504 GI:29351859
VERSION CB518504.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (Bases 1 to 631)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1..631

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE: 6837653"

/tissue_type="Whole brain"

/dev_stage="1, 5, and 15 days newborn"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH BMAP_GHO"

/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;

Site 2: Not I; the library was constructed according

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

with EcoR I adaptor, digested with NotI and then cloned

directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

is CGAACTGAAT. This library was created for the University

FEATURES

Location/Qualifiers

1..631

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE: 6837653"

/tissue_type="Whole brain"

/dev_stage="1, 5, and 15 days newborn"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH BMAP_GHO"

/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;

Site 2: Not I; the library was constructed according

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

with EcoR I adaptor, digested with NotI and then cloned

directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

is CGAACTGAAT. This library was created for the University

Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 151 a 183 c 180 g 117 t
ORIGIN

Alignment Scores:
Pred. No.: 1,37e-64 Length: 631
Score: 1098.00 Matches: 208
Percent Similarity: 99.52% Conservative: 0
Best Local Similarity: 99.52% Mismatches: 1
Query Match: 16.63% Indels: 0
DB: 14 Gaps: 0

US-09-754-997A-2 (1-1252) x CB518504 (1-631)

QY 470 GlyValAspAsnValGluTyrGlnPheAlaValAsnAsnAspThrThrGluLeuGlnVal 489
DB 3 CGAGTGGACAATGTGGAGTACCAGTTTGCAGTAAACATGACACACAGAGCTGCAGGTT 62
QY 490 ArgAspLeuGluProAsnThrAspTyrGluPheTyrValValAlaTyrSerGlnLeuGly 509
DB 63 CGGGACCTGGAAACCCCAACCGGATTATGAGTTCTACGTGGTGGCTTACTCCAGCTGGGG 122
QY 510 AlaSerArgThrSerProAlaLeuValHisThrLeuAspAspValProSerAlaAla 529
DB 123 GCAGCGGAACTCCAGCCGAGCCCTGGTGCATACATGACGATGTCCTCCAGCGGAGCA 182
QY 530 ProGlnLeuThrLeuSerProAsnProSerAspIleArgValAlaTyrLeuProLeu 549
DB 183 CCCAGCTTACCTTGTCCAGCCCAACCCCTCGGACATCAGGTGGCATGGCTGCCCTG 242
QY 550 ProSerSerLeuSerAsnGlyGlnValLeuLeuTyrTyrIleGluTyrGlyLeuGlyLys 569
DB 243 CCTCCAGCTGAGCAATGACAGGTGCTGAAGTACAAGATAGAGTACGTTTGGGGAAG 302
QY 570 GluAspGlnValPheSerThrGluValProGlyAsnGluThrGlnLeuThrLeuAsnSer 589
DB 303 GAAGATCAGTTTCTCCACCGAGGTGCTTGGAAATGAGACACACTTACGTTAAACTCA 362
QY 590 LeuGlnProAsnLysValTyrArgValArgIleSerAlaGlyThrGlyAlaGlyTyrGly 609
DB 363 CTTACGCCAACCAAGATGTACCGAGTCCGGATTTACGTGGCACTGGCTGGCTATGGA 422
QY 610 ValProSerGlnTrpMetGlnHisArgThrProGlyValHisAsnGlnSerHisValPro 629
DB 423 GTCCCTTCTCAGTGGATGACAGCAGGACACCTGGTGTGCACAAACAGAGCCATGTTCC 482
QY 630 PheAlaProAlaGluLeuLysValArgAlaLysMetGluSerLeuValValSerTrpGln 649
DB 483 TTTGGCCCTGCAGATTTGAAGGTGGAGGCAAGATGGAGTCCCTGGTGGTGTATGGCAG 542
QY 650 ProProHisProThrGlnIleSerGlyTyrLysLeuTyrTrpGlyGluValGlyThr 669
DB 543 CCGCCCTCCACCCACCCAGATCTCTGGATACAACTCTACTGGAGAGAGGTGGGACA 602
QY 670 GluGluAlaAlaAspGlyAspArgPro 678
DB 603 GAGGAGGAGGAGATGGTACCGCCCC 629

RESULT 11

CB988160
LOCUS
DEFINITION
AGENCY 13892946 NIH MGC 147 Homo sapiens cDNA clone
IMAGE:30346067 5', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 759)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM384 row: p column: 12
High quality sequence stop: 606.
Location/Qualifiers

FEATURES
source

1..759
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30346067"
/tissue_type="Human Placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH MGC 147"
/note="Organ: placenta; Vector: pBluescriptR; Site 1:
all-XhoI; Site 2: BamH; Oligo-dT primed using primer
5'-TTTTTTTTTTTTTNN-3', size-selected for average
insert size 2.3 kb and normalized to 10^5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: This is
a NIH MGC library."

BASE COUNT 161 a 229 c 237 g 131 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 2,02e-64 Length: 759
Score: 1097.00 Matches: 218
Percent Similarity: 94.17% Conservative: 8
Best Local Similarity: 90.83% Mismatches: 11
Query Match: 16.61% Indels: 3
DB: 14 Gaps: 1

US-09-754-997A-2 (1-1252) x CB988160 (1-759)

QY 378 GlyArgValLysValGlnGlyGlySerLeuValIleThrGlnIleGlyLeuGln 397
DB 27 GGGCGGTCAAGTCCAGGGCGGTGGCAGCTGGTCATCACAGATCGGCTGCAG 86
QY 398 AspAlaGlyTyrTyrGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAla 417
DB 87 GACGCGGTACTTACAGTGGTGGCTGAGAACAGCGGGGAATGGCGTGGCGG 146
QY 418 ProLeuAlaValValArgGluGlyLeuProSerAlaProThrArgValThrAlaThr 437
DB 147 TCGCTGGCGTGGTGGCGAGGGGCTGCCAGCCCCCAGCGGGTCACTGCTAG 206
QY 438 ProLeuSerSerSerValLeuValAlaTyrGluArgProGluLeuHisSerGluGln 457
DB 207 CCAGTGGAGAGTCCGCTGTGTGGTGGCGGGCGGCGAGATGCACAGCGAGCAG 266
QY 458 IleIleGlyPheSerLeuHisTyrGlnLysAlaArgGlyValAspAsnValGluTyrGln 477
DB 267 ATCATCGGCTTCTCTCCACTACCAAGAGGACAGCGGGCATGCACAAATGTGAATACCA 326
QY 478 PheAlaValAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAsp 497
DB 327 TTTGAGTGAACAACGACACCACTACAGGTTCGGGATCGGAGCCCAACACAGAT 386
QY 498 TyrGluPheTyrValValAlaTyrSerGlnLeuGlyAlaSerArgThrSerProAla 517

Db 387 TATGAGTTCTACGTGGCTACTCCAGCTGGGAGCCAGCCAGCTCCACCCAGCA 446

Qy 518 LeuValHisThrLeuAspValProSerAlaProGlnLeuThrLeuSerSerPro 537

Db 447 CTGGTGACACACTGGATGTCCCACTAGTCAGCAGCCAGCTCTCCCTGTCAGCC 506

Qy 538 AsnProSerAspIleArgValAlaThrLeuProLeuProSerSerLeuSerAsnGlyGln 557

Db 507 AACCTTCGGACATCAGGGTGGCGTGGCTGCTCCCTCCAGCTGAGCAATGGGGAG 566

Qy 558 ValLeuLysTrpLysIleGluTyrGlyLeuGlyLysGluAspGlnValPheSerThrGlu 577

Db 567 GTGGTGAAGTACAAGATAGATACGGTTTGGGAAGGAAGATCAGATTCTCTACTAG 626

Qy 578 ValProGlyAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyrArg 597

Db 627 GTGGGAGGAATGACACACAGCTTATGCTGAACCTCGCTTACGCCAACAAGGTGTATCGA 686

Qy 598 ValArgIle-SerAlaGlyThrGly-----AlaGlyTyrGlyValProSerGlnTrp 614

Db 687 GTACGGATTTTCGGCTGGTACAGCAGCCCGGGCTTCNGGGGCCCCCTCCAGGTGG 744

RESULT 12

LOCUS BY712424

DEFINITION BY712424 RIKEN full-length enriched, 10, 11 days embryo whole body

ACCESSION BY712424

VERSION BY712424.1 GI:27123733

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 703)

AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chochia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, M.J., Perce, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempke, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yangisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kondo, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, M., Imotani, K., Iehi, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

MEDLINE 22354583

PUBMED 12466851

COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Teurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K., Iehi, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

1..703

Location/Qualifiers

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="2810488H22"

/tissue_type="whole body"

/dev_stage="10, 11 days embryo"

/clone_lib="RIKEN full-length enriched, 10, 11 days embryo whole body"

BASE COUNT 164 a 185 c 218 g 134 t 2 others

ORIGIN

Alignment Scores:

Pred. No.: 1,75e-63 Length: 703

Score: 1082.50 Matches: 210

Percent Similarity: 86.53% Conservative: 2

Best Local Similarity: 85.71% Mismatches: 3

Query Match: 16.39% Indels: 30

DB: 14 Gaps: 1

US-09-754-997A-2 (1-1252) x BY712424 (1-703)

Qy 497 AspTyrGluPheTyrValValAlaTyrSerGlnLeuGlyAlaSerArgThrSerSerPro 516

Db 2 GATTATGAGTTCTACGTGGTGGCTACTCCAGCTGGGGCCAGCCAGCTCCAGCCCA 61

Qy 517 AlalaValHisThrLeuAspValProSerAlaAlaProGlnLeuThrLeuSerSer 536

Db 62 GCCCTGGTGATACATCTGGAGATGTCCCGAGCCAGCAGCCAGCTTACCTTGTCCAGC 121

Qy 537 ProAsnProSerAspIleArgValAlaThrLeuProLeuProSerSerLeuSerAsnGly 556

Db 122 CCCAACCCCTCGGACATAGAGGGTGGCATGTGCTCCCTGCTCCAGCTGAGCAATGGA 181

Qy 557 GlnValLeuLysTrpLysIleGluTyrGlyLeuGlyLysGluAspGlnValPheSerThr 576

Db 182 CAGGTGCTGAAGTACAAGATAGATACGGTTTGGGAAGGAAGATCAGATTCTCTCACC 241

Qy 577 GluValProGlyAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyr 596

Db 242 GAGGTGCTGGAATGAGACACAACTTACGTGTTAACTCCTTACGCCAACAAGGTGAC 301

Qy 597 ArgValArgIleSerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGln 616

```

Db      302  CGAGTCCGGAATTCAGCTGCAGCTGGCGTGGCTATGAGTCCCTTCTCAGTGGATGCAG 361
QY      617  HsAtrqThrProGlyValHisAenGlnSerHis-ValProPheAlaProAlaGluLeuLy 636
Db      362  CACAGACACCTGTGTGCAACACGAGGCCATGATC----- 398
QY      636  sValAArgAlaLysMetGluSerLeuValValSerTrpGlnProProHisProThrG 656
Db      398  ----- 398
QY      656  nIleSerGlyTyrLysLeuTyrTrpGlyGluValGlyThrGluGluGluAlaAspGlyAs 676
Db      399  ----TCTGGATACAACTCTACTGGAGAGAGGTGGCAACAGAGGAGGAGGCGCATGTGCA 454
QY      676  pArgProGlyGlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLy 696
Db      455  CCGCCCCCAGGGGGTGGTGGAGATCAAGCTTGGAGCTGGGGCCCGTGGCGCTGAAGAA 514
QY      696  sLysValLysGlnTyrGluLeuThrGlnLeuValProGlyArgProTyrGluValLysLe 716
Db      515  GAAAGTAGGACGATGATGACTGACCCAGTTAGTCCCTGGCAGGCTGTACGAGGTGAAGCT 574
QY      716  uValAlaPheAsnLysHisGluAspGlyTyrAlaAlaValTrpLysGlyLysThrGluLy 736
Db      575  CGTAGCTTTCACANACACAGGACGCTACGCTGTGTGGAAGGCGAAGACGGAGAA 634
QY      736  sAlaProThrPro 740
Db      635  GCGGCCCGCCGCA 647

```

RESULT 13

```

BU057120
LOCUS      591 bp mRNA linear EST 26-AUG-2002
DEFINITION UI-M-FOO-caa-e-21-0-UI.r1 NIH_BMAP_F00 Mus musculus cDNA clone
IMAGE:6408572 5', mRNA sequence.

```

```

BU057120
VERSION    BU057120.1 GI:22497197

```

```

KEYWORDS   Mus musculus (house mouse)
SOURCE     Mus musculus

```

ORGANISM

```

REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS    1 (bases 1 to 591)
            NIH-MGC http://mgc.nci.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov

```

```

Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

```

```

Seq primer: pyx-5.

```

FEATURES

source

```

1..591
    Location/Qualifiers
        /organism="Mus musculus"
        /mol_type="mRNA"
        /strain="C57BL/6"
        /db_xref="taxon:10090"
        /clone="IMAGE:6408572"
        /tissue_type="whole brain"
        /dev_stage="embryo 12.5dpc"
        /lab_host="PH108 (T1 phage resistant)"
        /clone_lib="NIH_BMAP_F00"
        /note="Organ: Brain; Vector: pyx- Asc; Site 1: EcoR I;
            Site 2: Not I; The library was constructed according
            Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

```

1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pyX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TCAGAGAGCC. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 126 a 179 c 179 g 106 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 3,55e-61 Length: 591
Score: 1047.00 Matches: 193
Percent Similarity: 98.98% Conservative: 1
Best Local Similarity: 98.47% Mismatches: 2
Query Match: 15.85% Indels: 0
DB: 13 Gaps: 0

US-09-754-997A-2 (1-1252) x BU057120 (1-591)

```

QY      1005  ArgLeuGlyProProSerValProAlaAlaHisGluLeuGluSerLeuValHisProArg 1024
Db      3  CGACTTGGGCCCCCAGTGTCCCTGCTGCCCATGAGTTGGAGTCCCTGTGTCATCTCTGT 62
QY      1025  ProGlnAspTrpSerProProSerAspValGluAspLysAlaGluValHisSerLeu 1044
Db      63  CCCCAGGATTGGTCCCCACCCCTCAGATGTGGAGACACAGGCTGAAGTACACAGCCTT 122
QY      1045  MetGlyGlySerValSerAspCysArgGlyHisSerLysArgLysIleSerTrpAlaGln 1064
Db      123  ATGGGTGGCAGTGTTCAGATTGCGGGGCCACTCCAGAGAGAAGATCTCTCTGGGCTCAG 182
QY      1065  AlaGlyGlyProAsnTrpAlaGlySerTrpAlaGlyCysGluLeuProGlnGlySerGly 1084
Db      183  GCAGGGGGGCCAAACTGGGCGAGGCTCTTGGGCGAGGCTGTGAGCTGCCCCCAGGGTAGTGGT 242
QY      1085  ProArgProAlaLeuThrArgAlaLeuLeuProProAlaGlyThrGlyGlnThrLeuLeu 1104
Db      243  CCAAGGCGCGCTCTGACCGCTGCTCTGCTGCTCCAGCGGGAACCGGCGCAGACACTGCTG 302
QY      1105  LeuGlnAlaLeuValTyrAspGlyIleLysSerAsnGlyArgLysLeuProSerProAla 1124
Db      303  CTGCAAGCCCTGGTATATGACGCCCATNAAGAGCAACGGGAGAAAGAGCCGTCGCCAGCC 362
QY      1125  CysArgAsnGlnValGluAlaGluValIleValHisSerAspPheGlyAlaSerLysGly 1144
Db      363  TGCAGGATTCAGTGGAGCTGAGGTGTCATTGTTCCTCGACTTCGATTCGATCCAAAGA 422
QY      1145  CysProAspLeuHisLeuGlnAspLeuGluProGluGluProLeuThrAlaGluThrLeu 1164
Db      423  TGTCTCTGACCTCCACCTCCAGACCTCGAGCCAGAGGAACCACTGACTGCAGAGACTCTG 482
QY      1165  ProSerThrSerGlyAlaValAspLeuSerGlnGlyAlaAspTrpLeuGlyArgGluLeu 1184
Db      483  CCTTCCAGCTCTGGAGCTGTGGATCTGTCTCAAGAGAGCANACTGGCTGGGCAAGGAGCTG 542
QY      1185  GlyGlyCysGlnProThrThrSerGlyProGluArgLeuThrCysLeu 1200
Db      543  GGAGGGGTGCCAACCAACCAACCCAGTGGGCGCAGAGAGGCTCACCTGCTTGT 590

```

RESULT 14

AK051027

LOCUS

DEFINITION

AK051027

ACCESSION

VERSION

AK051027 2906 bp mRNA linear HTC 05-DEC-2002
Mus musculus 9 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:D030056K15 product:putative neuronal cell
adhesion molecule, full insert sequence.
AK051027
AK051027.1 GI:26341679

KEYWORDS	HTC; CAP trapper.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
REFERENCE	
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, K. I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischnann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Mateu, Y., Nikaudo, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bona, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. P., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409 (6821), 685-690 (2001)
MEDLINE	21085660
PUBMED	11217851
REFERENCE	
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
MEDLINE	12127851
PUBMED	12127851
REFERENCE	
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurahara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, K., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
TITLE	Alignment Scores:
JOURNAL	2.97e-60 Length: 2906
MEDLINE	1046.00 Matches: 328
PUBMED	Percent Similarity: 38.49% Conservative: 135
REFERENCE	Best Local Similarity: 27.27% Mismatches: 307
AUTHORS	Query Match: 15.84% Indels: 434
TITLE	DB: 11 Gaps: 34
JOURNAL	US-09-754-997A-2 (1-1252) x AK051027 (1-2906)
MEDLINE	24 LeuProLeuProGlnGluThr-----ThrVallyLeuSerCysAspGlu 38
PUBMED	342 TTGCGCTGCTGCGCCAGCTGGTCTGGGCCATTCTGCTGCACTTTTCTGTG 401
REFERENCE	39 GlyProLeuGlnValIleLeuGlyProGluGlnAlaValValLeuAspCysThrLeuGly 58
AUTHORS	402 GAGCCAAATGATGATGATATTGCCAACCTGGCGCAGCCCATAGTGTCTGGCTGCAAGTGGAG 461
TITLE	59 AlaThrAlaAlaGlyProProThrArgValThrTrpSerLysAspGlyAspThrValLeu 78

Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/
 Location/Qualifiers
 1. 2906
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM_DB:D030056K15"
 /db_xref="taxon:10090"
 /clone="D030056K15"
 /tissue_type="whole body"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="9 days embryo"
 228..2669
 /note="unnamed protein product; putative putative neuronal cell adhesion molecule (MGD|MG1:1202390, GB|NM_008988, evidence: BLASTN, 100%, match=726)"
 /codon_start=1
 /protein_id="BAC34502.1"
 /db_xref="GI:26341680"
 /translation="MAEPTASPRRLPALRRPGLPPLPPPPPLLLLLLLPLPAP SLGLHSAALAFSVEPNDDIANPGQIVLCKVEGTPPVQVQVNRKGAELPEGTHTTL LANGSLIHIFRLLEGSGSPDEGDKVCAQNRGLVSRKARLQAATMSDPVHPQAV TGEGGVARQCQIHGPKLITWKRNVFIDTDDERTLLPKGLVLTGLRADSGI PHCVASNIASVRVSHGARLTVSGSGVTEPTLVGPENLTLTVHGTAVLVCATGQ PRFTVSWDLRGPVIGEGVLGTNLIIISDVTVHSGVYVCAANRPGTRVRAAQ RLNVQAPAEVQHPQISIRPAGTAMFTCAQGEPPPHVTWLNQVGLGAGHVLKN NNSTLSISGVGPDEAIYQCAENIAGSSQASARLTVLWAGLPGPPNVRVASVST EVVSWSEPLAHTKEILIGYVHLHKAADSPLKYQEAIVSKSTFOHLYRDLPESTAYSP YIKAYTRGASLASVPTLSTLGEAPVPPPLSVLLGSSSLQLLWKPRLAQNGGF KLFRPVSAVSFTGPILLPTVSSNLSQDPSTVYEVKLLAYNHQDGNATVRPVS LKGSERALTTPCDCKEDVTNHTSTTGVIGIHGVTCTFCVLLFLGQRGLVLLC KDVENQLSPQGRSORDPGLALNGLSRGEGQLSRDKPVDKLEQLPPTAGSAA QPQSTPTDPAAPACBETLSMVLQGLFNLVAGRTTREATSPCAGPGVPVPPQDIGPVP LSEQQTQPPVAPAQ"

BASE COUNT 582 a 937 c 835 g 552 t
 ORIGIN
 Alignment Scores:
 Pred No.: 2.97e-60 Length: 2906
 Score: 1046.00 Matches: 328
 Percent Similarity: 38.49% Conservative: 135
 Best Local Similarity: 27.27% Mismatches: 307
 Query Match: 15.84% Indels: 434
 DB: 11 Gaps: 34
 US-09-754-997A-2 (1-1252) x AK051027 (1-2906)
 Qy 24 LeuProLeuProGlnGluThr-----ThrVallyLeuSerCysAspGlu 38
 Db 342 TTGCGCTGCTGCGCCAGCTGGTCTGGGCCATTCTGCTGCACTTTTCTGTG 401
 Qy 39 GlyProLeuGlnValIleLeuGlyProGluGlnAlaValValLeuAspCysThrLeuGly 58
 Db 402 GAGCCAAATGATGATGATATTGCCAACCTGGCGCAGCCCATAGTGTCTGGCTGCAAGTGGAG 461
 Qy 59 AlaThrAlaAlaGlyProProThrArgValThrTrpSerLysAspGlyAspThrValLeu 78

```
Db 462 GGCACC-----CTCCAGTCAAGTCTCTCTGGAGGAAGAAATGGAGCAGAGCTTACCA 512
Qy 79 GluHieGluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSer----ProLeu 97
Db 513 GAGGGACCCACACACCACTCTGCTGGCCAAATGGCTCTCTGATCACCACATTCACGGCTG 572
Qy 98 GluGlnGluAspSerAspGluGluAlaLeuArgIleTrpLysValThrGluGlySer 117
Db 573 GAGCAGGAGGTAGCCCTTCAGAT-----GAGGGTGAC 605
Qy 118 TyrSerCysLeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValLys 137
Db 606 TATGAATGTGGCTCAGAACCGCTTGGGCTGCTGGCTCAGTCGGAAGGCTCTCTCCAG 665
Qy 138 LeuAlaThrLeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGly 157
Db 666 GCTGCAACTATGTCCGACTTCCAGCTACACCCCGCCGCTCAGGGTGAGGAGGTGGT 725
Qy 158 ThrAlaArgPheGluCysHisThrLysGlyLeuProAlaProIleThrTrpGluLys 177
Db 726 GTGGCCCGCTTCCAGTGCCAAATCCATGGGCTTCCGAGCCCTGATCACAATGGGAGAG 785
Qy 178 AspGlnValThrVal---ProGluGluProArgLeuIleThrLeuProLysTrpLeuLeu 196
Db 786 AACAGAGTTCCTCCATTGACAGCGAGATGAGAGGTACACACTACTGCCAAGGAGTCCTG 845
Qy 197 GlnIleLeuAspValGlnAspSerAspAlaGlySerTyrArgCysValAlaThrAsnSer 216
Db 846 CAGATAACAGGACTTCGGGCTGAGGACAGCGGCATCTTCCACTGTGTGGCTCAAAACATT 905
Qy 217 AlaArgGlnArgPheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGlu 236
Db 906 GCTAGTGTGGGTGACGCCACGGGGCCAGGCTCACTGTGTCA-----GGCTCGGGCTCA 959
Qy 237 AlaThrArgGlyGlnAspValValIleValAlaAlaProGluAsnThrThrValValSer 256
Db 960 GGGACCTTACAGGAGCCCACTTCTGTG---GGGCTTGAGAACCTCACCCCTGACAGTG 1016
Qy 257 GlyGlnAsnValValMetGluCysValAlaSerAlaAspProThrProPheValSerTrp 276
Db 1017 CACCAGACTCGGGTGTGGAGTGTGTGCCACGGGCAACCGCGCCCATTTGTGTCTGTG 1076
Qy 277 ValArgGlnAspGlyLysProIleSerThrAspValIle---ValLeuGlyArgThrAsn 295
Db 1077 AGCGCTGTGATGGCCGCCCATTTGGGGTGAGGGGCATCCAGGTGCTGGGCACAGGAAC 1136
Qy 296 LeuLeuIleAlaSerAlaGlnProArgHisSerGlyValTyrValCysArgAlaAsnLys 315
Db 1137 CTTATCATCTCAGATGTGACTGTCAGCATTCAGGTGTCTATGTCTGTGTGAGCCACAGG 1196
Qy 316 ProLeuThrArgAspPheAlaThrAlaAlaGluLeuArgValLeuAlaAlaProAla 335
Db 1197 CTTGGCACCCGGGTGAGAGGACGCGCGGTGCGGTGCTGTGTGTGTGTGTGTGTGTGTGT 1256
Qy 336 IleSerGlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCys 355
Db 1257 TTTGTTTCCAGTATCCAGTCCATCTCCAGGCTGCTGGGACCCAGCCATTTTCCACTGC 1316
Qy 356 ArgAlaSerGlyGluProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArg 375
Db 1317 CAAGCGGAGGGGAGCCACCGCTCATGTGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1376
Qy 376 ProAsnGlyArgValLysValGlnGlyGlySerLeuValIleThrGlnIleGly 395
Db 1377 GCAGGAGGCCATGTGAGGCTCAGAACCAATAACAGCACACTGAGCAATTTCTGGAGTCGA 1436
Qy 396 LeuGlnAspAlaGlyTyrTyrGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAla 415
Db 1437 CCTGAAGTAGGGCCATCTACCAATGTGTGGCTGAGAACATTTGTGGCTCGTCCGAAGCC 1496
Qy 416 AlaAlaProLeuAlaValValArgGluGlyLeuProSerAlaProThrArgValThr 435
Db 435
```

```
1497 AGCGCCAGGCTGACCGTGTGTGGCTGAGGGTTGCCAGGCGCCCCACGCAATGTCCGA 1556
Qy 436 AlaThrProLeuSerSerSerValLeuValAlaTrpGluArgProGluLeuHisSer 455
Db 1557 GCTGTCTGTGTCTCTCCACTGAGTTGAGTGTCTCTGAGTGTGAGCTCTGCGCCACACC 1616
Qy 456 GluGlnIleIleGlyPheSerLeuHisTyrGlnLysAlaArgGlyValAspAsnValGlu 475
Db 1617 AAAGAGATCATTTGGCTGTATGCTTTCACATCATCAGGAGGCTGCCGACTCACCCCAAGCTAGAG 1676
Qy 476 TyrGlnPheAlaValAsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsn 495
Db 1677 TACCAGGAGCTGTTCAGCAAGAGCACCTTTCAGCACCTGCTGAGAGAGCTTTCAGGCTCC 1736
Qy 496 ThrAspTyrGluPheTyrValValAlaTyrSerGlnLeuGlyAlaSerArgThrSerSer 515
Db 1737 ACTGCTCAGCTTCTTACATCAAGCCTACACACCGAGGGGCGCCAGTTTACCTCTGTCTC 1796
Qy 516 ProAlaLeuValHisThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSer 535
Db 1797 CCCACCTGGCCAGCACCTCTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1853
Qy 536 SerProAsnProSerAspIleArgValAlaTrpLeuProLeuProSerSerSerLeuSerAsn 555
Db 1854 TTGCTGGGCGAGCTCTCTTACACTGCTATGAGAGCCCTGCCCA-----1898
Qy 556 GlyGlnValLeuLysTyrLysIleGluTyrGlyLeuGlyLysGluAspGlnValPheSer 575
Db 1898 -----1898
Qy 576 ThrGluValProGlyAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysVal 595
Db 1898 -----1898
Qy 596 TyrArgValArgIleSerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMet 615
Db 1899 -----CGGCTGGCCCA-----1910
Qy 616 GlnHisArgThrProGlyValHisGlnSerHisValProPheAlaProAlaGluLeu 635
Db 1911 -----CACAA-----1916
Qy 636 LysValArgAlaLysMetGluSerLeuValValSerTrpGlnProProHisProThr 655
Db 1916 -----1916
Qy 656 GlnIleSerGlyTyrLysLeuTyrTrpGlyValGlyThrGluGluAlaAspGly 675
Db 1917 -----GGAGGCTTCAAGCTGTTTACCGCCAGTCAGTCAACCTCC-----1958
Qy 676 AspArgProGlyGlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLys 695
Db 1959 -----TTCACTGGCCCATCTTGTCTCCCT 1982
Qy 696 LysLysValLysGlnTyrGluLeuThrGlnLeuValProGlyArgProTyrGluValLys 715
Db 1983 GGGACTGTCTCTTACACCTCAGCTTACCCAGCTTGCACCCAGCAGCTGTACGAGGTGAAG 2042
Qy 716 LeuValAlaPheAsnLysHisGluAspGlyTyrAlaAlaValTrpLysGlyLysThrGlu 735
Db 2043 CTCTGGCTTACACCAACGATGAGATGGA-----2072
Qy 736 LysAlaProThrProAspLeuProIleGlnArgGlyProProLeuProProAlaHisVal 755
Db 2072 -----2072
Qy 756 HisAlaGluSerAsnSerThrSerIleTrpLeuArgTrpLysLysProAspPheThr 775
Db 2072 -----2072
Qy 776 ThrValLysIleValAsnTyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSer 795
Db 2073 -----AAGGCCACGGTCCGATTT-----GTGTCTCTGAGGGAGCTCT 2111
```

QY 796 LeuValThrTyrThrSerSerGlyGluAspIleLeuIleGlyGlyLeuValProPhe 815
 Db 2111 ----- 2111
 QY 816 ThrLysTyrGluPheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySer 835
 Db 2111 ----- 2111
 QY 836 ValValGluArgSerThrLeuProAspArgProSerThrProProSerAspLeuArgLeu 855
 Db 2112 ----- 2112
 QY 856 SerProLeuThrProSerThrValArgLeuHisTyrCysProProThrGluProAsnGly 875
 Db 2151 GAAGATGTCACC----- 2162
 QY 876 GluIleValGluTyrLeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTyr 895
 Db 2163 ----- 2171
 QY 896 ThrLeuLeuThrThrGluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAsp 915
 Db 2171 ----- 2171
 QY 916 ThrArgTyrPhePheLysMetGlyAlaArgThrGluValGlyProGlyProPheSerArg 935
 Db 2171 ----- 2171
 QY 936 LeuGlnAspValIleThrLeuGlnGluThrPheSerAspSerLeuAspValHisAlaVal 955
 Db 2172 ----- 2177
 QY 956 ThrGlyIleIleValGlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAla 975
 Db 2178 ACAGCATTTGATCGCATCCACATCGGAGTCACCTGCATCATCTTCTCGCTCTCTT 2237
 QY 976 GlyLeuArgGlnSerSerHisArgGluAlaLeuProGlyLeuSerSerGlyThrPro 995
 Db 2238 ----- 2269
 QY 996 -----GlyAsnProAlaLeuTyrThrArgAlaArgLeuGlyProProSerVal 1011
 Db 2270 GTGTAAGGATGTGGAAACACGCTCTC----- 2302
 QY 1012 ProAlaAlaHisGluLeuGluSerLeuValHisProArgProGlnAspTyrProProPro 1031
 Db 2303 CCAGGGTCCCGGAGCCAGCCAGCCCTGTATCTCTGGCCCT--GAATGAGACTGAGCCG 2359
 QY 1032 -----ProSerAspValGluAspLysAlaGluValHisSerLeuMetGlyGly 1047
 Db 2360 AGGAGAAGGGGGCCAGCTGAGCCGAGATGAAGCCGGTGGAGCCGACAGGAGTGGAC 2419
 QY 1048 SerValSerAspCysArgGlyHisSerLysArgLysIleSerTrpAlaGlnAla----- 1065
 Db 2420 GCTGTTTCCAC-----AGCTGGGTGAGCTGCACAGCC 2452
 QY 1066 -----GlyGlyProAsnTrpAla 1071
 Db 2453 AGGCTCACACCCACCGATCCTGAGCCCTGTCTCATGTGAGGAGCCACCGCTCTCCAT 2512
 QY 1072 GlySer----- 1073
 Db 2513 GGTTCAGCTTCAGGCTTCACTTGTGGCAGGAGGACACAGAGCCACGTCCTCCCTG 2572
 QY 1074 -----TrpAlaGlyCysGluLeuProGlnGlySerGlyProArgProAlaLeuThr 1090
 Db 2573 TGCAGGCGCTGGGCC-----TGTCACAGCAGCCACGAGATATAGCCCTGTCCCT 2623
 QY 1091 ArgAlaLeuLeuProProAlaGlyThrGlyGlnThrLeuLeuLeuGlnAlaLeuValTyr 1110
 Db 2624 CAGTGAAGGACAGACCGAGCTCCAGCAGTCGACGACCCCGAGTGCATCTGTGTCC 2683

QY 1111 AspGlyIleLysSerAsnGlyArgLysLysProSerProAlaCysArgAsnGlnValGlu 1130
 Db 2684 CGAGGGCGGTATCCCAT-----TTCAGGTCAA 2713
 QY 1131 AlaGluValIle---ValHisSerAspPheGly---AlaSerLysGlyCysProAspLeu 1148
 Db 2714 GCGAGGTTTCTCTCTGTTCTGGGATTTGGCGCTCCGACGTAAGA----- 2761
 QY 1149 HisLeuGlnAsp-LeuGluProGluGluProLeu-----ThrAlaGluThrLeuProSe 1166
 Db 2762 -----CAGGGTTCTCAACACACAGGACCGAGTGGTCTCTACAGCAGGATCTCCCTG 2815
 QY 1166 rThrSer 1168
 Db 2816 CTCGAGC 2822

RESULT 15

BB624536 651 bp mRNA linear EST 31-AUG-2001
 BB624536 RIKEN full-length enriched, 15 days embryo male testis Mus
 musculus cDNA clone 8030498014 5', mRNA sequence.
 BB624536
 BB624536.1 GI:15398322
 EST.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 651)
 Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
 Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,
 M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
 Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,
 D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
 Tagami,M., Tagawa,A., Takahashi,P., Takeda,Y., Tanaka,T., Toya,T.,
 Muramatsu,M. and Hayashizaki,Y.
 RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
 Unpublished

TITLE

JOURNAL
 COMMENT
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216

Email: genome-res@gsr.riken.go.jp
 URL: http://genome.gsc.riken.go.jp/
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
 M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
 Watanabe,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
 S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
 Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
 Y. and Hayashizaki,Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A., Aizawa
 K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K., Arakawa,T.,
 Ishii,Y. and Hayashizaki,Y.
 Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct.
 Func. Genomics 2 pre, L72-L86 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.
 e mouse tissues.

FEATURES

Location/Qualifiers

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 14, 2003, 06:11:04 ; Search time 511 Seconds
(without alignments)
6613.887 Million cell updates/sec

Title: US-09-754-997A-2

Perfect score: 6604

Sequence: 1 MARADTGRGLLVLTFCLLSA.....CPLTVSFLPAPVSSAQVP 1252

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09754997/runat_14102003_061059_9696/app_query.fasta_1.1415
-DB=N_Geneseq_19Jun03 -QPM=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US09754997 @CNG 1 1 0 @runat_14102003_061059_9696 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query	Length	DB	ID	Description
1	6604	100.0	6176	22	AAD10021		Mouse Nope (neighb
2	5746	87.0	3753	22	AAI67204		Nucleotide sequenc
3	5610	84.9	3741	24	ABK11101		DNA encoding human
4	4903	74.2	2796	22	AAD10022		Mouse Nope (neighb
5	4361	66.0	3450	24	ABZ11219		Human polynucleoti
6	3167	48.0	2223	22	AAI67203		Nucleotide sequenc
7	1747	26.5	3453	25	ABX71435		Human cDNA encodin
8	1746	26.4	3210	25	ABX71432		Human cDNA encodin
9	1691	25.6	3874	25	ABX71447		Human cDNA encodin
10	1691	25.6	2976	25	ABX71437		Human cDNA encodin
11	1691	25.6	3219	25	ABX71440		Human cDNA encodin
12	1472	22.3	825	22	AAD10023		Mouse Nope (neighb
13	1458	22.1	3330	23	AA576658		DNA encoding novel
14	1384	21.0	2715	25	ABX71434		Human cDNA encodin
15	1384	21.0	2958	25	ABX71436		Human cDNA encodin
16	1339	20.1	2481	25	ABX71439		Human cDNA encodin
17	1339	20.1	2724	25	ABX71441		Human cDNA encodin
18	1144	17.3	2139	25	ABX71442		Human cDNA encodin
19	1144	17.3	2382	25	ABX71445		Human cDNA encodin
20	1067	16.2	2700	24	AAD36777		Human immunoglobul
21	958	14.5	4650	23	AA570835		DNA encoding novel
22	955	14.5	5553	24	AB576362		DNA encoding human
23	953	14.4	4820	25	ABT17038		Human MP21 gene ne
24	951	14.4	5297	24	ABV94279		Breast carcinoma r
25	951	14.4	5297	25	ABX10384		DNA encoding prote
26	943	14.3	4344	22	AAC90961		Human UNC-40 nucle
27	943	14.3	4608	16	AAQ80196		Deleted in colorec
28	943	14.3	4608	20	AAZ23431		Human DCC DNA. Ho
29	943	14.3	4608	25	ABT17039		Human MP21 gene DC
30	943	14.3	5186	12	AAQ12752		Deleted in Colorec
31	933	14.1	2695	23	AA576339		DNA encoding novel
32	909	13.8	1875	25	ABX71443		Human cDNA encodin
33	829	12.6	4248	22	AAC90960		C. elegans UNC-40
34	781	11.8	1644	25	ABX71444		Human cDNA encodin
35	781	11.8	1887	25	ABX71446		Human cDNA encodin
36	768	11.6	6279	23	ABU14245		Drosophila melanog
37	698	10.6	2450	23	AA576657		DNA encoding novel
38	667	10.1	7630	23	ABL30223		Drosophila melanog
39	648	9.8	6734	24	AB199344		Mouse ischaemic co
40	647	9.8	1143	25	ABX71433		Human cDNA encodin
41	638	9.7	6000	16	AAQ86478		Human FTP-OB. Hom
42	638	9.7	6000	18	AA785389		Human protein tyro
43	638	9.7	6000	20	AA506095		Human protein tyro
44	638	9.7	6000	22	AA502144		Human protein tyro
45	637	9.7	6374	24	AAD36781		Human immunoglobul

ALIGNMENTS

RESULT 1
AAD10021
ID AAD10021 standard; cDNA; 6176 BP.
XX
XX AAD10021;
XX
XX 12-SBP-2001 (first entry)
XX
XX Mouse Nope (neighbour of punc ell) cDNA.

XX Mouse; Nope; neighbour of punc ell; cytostatic; neuroprotective; vaccine;
XX gene therapy; cerebroprotective; colonic cancer; mental retardation;
XX tumour suppressor; chromosome 9; transgenic animal; genetic disorder;
XX obesity; Bardet-Biedl syndrome; autosomal recessive disorder; retinitis;
XX polydactyly; pigmentosa; hypogonadism; hypertension; diabetes mellitus;
XX renal anomaly; cardiovascular anomaly; ss.

OS Mus musculus.
 PH Key
 FT CDS
 FT Location/Qualifiers
 FT 1..3759
 FT /tag= a
 FT /product= "Mouse Nope protein"
 FT 1..63
 FT /tag= b
 FT 64..3756
 FT /tag= c
 FT /product= "Mature mouse Nope protein"
 FT 1..3756
 FT /tag= d
 FT /note= "This region is specifically referred as
 FT SEQ ID NO: 45 in the specification"
 FT 64..2859
 FT /tag= e
 FT /note= "Extracellular domain"
 FT 139..381
 FT /tag= f
 FT /label= Ig1
 FT /note= "Immunoglobulin domain 1"
 FT 463..654
 FT /tag= g
 FT /label= Ig2
 FT /note= "Immunoglobulin domain 2"
 FT 766..954
 FT /tag= h
 FT /label= Ig3
 FT /note= "Immunoglobulin domain 3"
 FT 1039..1233
 FT /tag= i
 FT /label= Ig4
 FT /note= "Immunoglobulin domain 4"
 FT 1285..1533
 FT /tag= j
 FT /label= Fn1
 FT /note= "Fibronectin-type III domain 1"
 FT 1579..1827
 FT /tag= k
 FT /label= Fn2
 FT /note= "Fibronectin-type III domain 2"
 FT 1888..2175
 FT /tag= l
 FT /label= Fn3
 FT /note= "Fibronectin-type III domain 3"
 FT 2248..2493
 FT /tag= m
 FT /label= Fn4
 FT /note= "Fibronectin-type III domain 4"
 FT 2542..2793
 FT /tag= n
 FT /label= Fn5
 FT /note= "Fibronectin-type III domain 5"
 FT 2932..3756
 FT /tag= o
 FT /note= "Intracellular domain"
 FT 3819..5682
 FT /tag= p
 FT 3'UTR
 FT WO200149714-A2.
 PN 12-JUL-2001.
 PD 26-OCT-2000; 2000WO-US29698.
 PP 04-JAN-2000; 2000US-0174496.
 PR 19-MAY-2000; 2000US-0205789.
 PR (NEUR-) NEUROSCIENCES RES FOUND INC.
 PA Salbaum JM;
 PI
 XX

DR WPI; 2001-441846/47.
 DR P-PSDB; AAE05251.
 XX Murine Nope polypeptides and nucleic acids useful for preventing, -
 FT diagnosing and treating colonic cancer and Bardet-Biedl syndrome -
 XX Claim 10; Fig 2A; 99pp; English.
 XX The present invention relates to Nope (neighbour of punc ell) which is
 CC used in the prevention, treatment and diagnosis of diseases associated
 CC with inappropriate Nope expression such as cancers especially colonic
 CC cancer and genetic disorders, as Nope is thought to be a tumour
 CC suppressor. Nope gene is located on chromosome 9 and is used in gene
 CC therapy. Nope is used as vaccine. Nope gene may be administered to treat
 CC diseases by rectifying mutations or deletions in a patient's genome that
 CC affect the activity of Nope by expressing inactive proteins or to
 CC supplement the patients own production of Nope polypeptides. Nope gene
 CC is used to study the expression and function of Nope polypeptides and
 CC their role in metabolism through the creation of transgenic animal
 CC models. The anti-Nope antibodies and Nope antagonists may also be used
 CC to down regulate Nope expression and activity for the treatment of
 CC Bardet-Biedl syndrome which is an autosomal recessive disorder
 CC characterised by mental retardation, obesity, polydactyly, retinitis
 CC pigmentosa and hypogonadism. Patients with Bardet-Biedl syndrome have a
 CC high incidence of hypertension, diabetes mellitus and renal and
 CC cardiovascular anomalies. The present sequence is mouse Nope (neighbour
 CC of punc ell) cDNA.
 XX SQ Sequence 6176 BP; 1316 A; 1756 C; 1634 G; 1470 T; 0 other;

Alignment Scores:
 Pred. No.: 1.28e-307 Length: 6176
 Score: 6604.00 Matches: 1252
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0
 US-09-754-997A-2 (1-1252) x AAD10021 (1-6176)

Qy 1 MetalArgAlaAspThrGlyArgGlyLeuValLeuThrPheCysLeuLeuSerAla 20
 Db 1 ATGGCGGGGGGGGACACGGGGCGGGCTCTGGTGTGCTGACCTTTCTGCTGTGTCGCG 60
 Qy 21 ArgGlyGluLeuProLeuProGlnGluThrThrVallyLeuSerCysAspGluGlyPro 40
 Db 61 CGCGGGGAGCTGCCATTGCCCCAGGACAACTGTCAAGCTGAGCTGTGATGAGGAGCC 120
 Qy 41 LeuGlnValIleLeuGlyProGluGlnAlaValValLeuAspCysThrLeuGlyAlaThr 60
 Db 121 CTGCAGTGATCTCTGGGCTGTGAGCAGGCTGTGGTGTGCTGACCTTTGGGGGCTACA 180
 Qy 61 AlaAlaGlyProProThrArgValThrTrpSerIysAspGlyAspThrValLeuGluHis 80
 Db 181 GCTGTGGGCTCCGACCCAGGCTGACATGAGGAGCAAGATGAGAGACACTGTACTAGAGCAT 240
 Qy 81 GluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGluGlnGlu 100
 Db 241 GAGAACCTGCACCTGCTACCCAAATGGCTCTCTGTGGCTGTCTCTCACCCCTAGAGCAAGAA 300
 Qy 101 AspSerAspAspGluGluAlaLeuArgIleTrpIysValThrGluGlySerTySerCys 120
 Db 301 GACAGCGATGATGAGGAGCTCTTAGGATCTGGAGAGTCACTAGGGGAGCTATTCTCTGT 360
 Qy 121 LeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValValValLeuAlaThr 140
 Db 361 CTGGCCACACAGCCCGCTAGGAGTGTGGCCAGCCAGGTTGCTGTGTCAAGCTTGCCACA 420
 Qy 141 LeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThrAlaArg 160
 Db 421 CTCGAAGACTTCTCTCTGCACCCGAGTCCAGATTTGGAGGAGAGAACGGGACGACGC 480
 Qy 161 PheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAspGlnVal 180

Db 481 TTTGAATGCCACACAGGGGCTTCCAGCCCCCAATTAATCTTGGGAAAGGACAGGTTG 540
Qy 181 ThrValProGluGluProArgLeuLeuThrLeuProLysTrpLeuLeuGlnIleLeuAsp 200
Db 541 ACCGTGCTCAGGAGCCCGGCTCATCTCTTCCCAAGTGGCTCTCTCCAGATCTAGAT 600
Qy 201 ValGlnAspSerAspAlaGlySerTyrArgCysValAlaThrAsnSerAlaArgGlnArg 220
Db 601 GTCCAGACAGTAGTAGGAGCTCTACCGCTGCGTGGCCACCAATTCAGCCCGCCACCA 660
Qy 221 PheSerGlnGlnAlaSerLeuThrValAlaLeuArgGlySerLeuGlnAlaThrArgGly 240
Db 661 TTCAGCCAGAGGGCTCGCTCACTGTGGCCCTCAGAGGGCTTTTGGAGGGCTTACCGAGGG 720
Qy 241 GlnAspValValIleValAlaAlaProGluAsnThrThrValValSerGlyGlnAsnVal 260
Db 721 CAGGATGTGTTCATTGTGGGAGCCCGGAGAACACACACAGGTAGTGTCTGGACAGATGTA 780
Qy 261 ValMetGlnCysValAlaSerAlaAspProThrProPheValSerTrpValArgGlnAsp 280
Db 781 GTGATGGAGTGGCGGCTCTGTCTGACCCCACTTCTGTGTCTGGTCCGACAGGAT 840
Qy 281 GlyLysProIleSerThrAspValIleValIleValLeuGlyArgThrAsnLeuLeuIleAs 300
Db 841 GGAAGGCTATCTCCACGGATGTCATGTCTGGGCGGACCAATCTACTCATGCGCCAGC 900
Qy 301 AlaGlnProArgHisSerGlyValTyrValCysArgAlaAsnLysProLeuThrArgAsp 320
Db 901 GCGCAGGCTCGGCATCTCGAGTCTATGTCTGCGAGCCACCAAGCCCTCTACGCGTGAC 960
Qy 321 PheAlaThrAlaAlaAlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaPro 340
Db 961 TTGCGCACTCGCGCTGTGAGTCCGAGTGTCTGCTGCCCGGAGCCATCTCGCAGGCACC 1020
Qy 341 GluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGlu 360
Db 1021 GAGCGCTCTCGCGAGCGCGGCGAGCACCGCGCGCTTCTGTGTCCGGGCGTCCGGGAG 1080
Qy 361 ProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgProAsnGlyVal 380
Db 1081 CCAGCGCGCGGCTGCACTGGCTGCACGAGGGATCCCGTGTGACCAATGGCGCGCTC 1140
Qy 381 LysValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGly 400
Db 1141 AAGGTGCAGGCGGTGGCGGAGCTTGTCTCATCTCAGATCGCGCTCGCAGGAGCGCTGC 1200
Qy 401 TyrTyrGlnCysValAlaGlnAsnSerAlaGlyThrAlaCysAlaAlaProLeuAla 420
Db 1201 TACTACCAAGTGCGTAGCAGAAAACAGCGCGGAACTGCCTGTGCGCGTGGCGCCCTGG 1260
Qy 421 ValValValArgGluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSer 440
Db 1261 GTAGTGTGCGGAGGGGCTGCCGAGCGCCCGGACTCGGGTCAAGCCAGCGCGCTGAGC 1320
Qy 441 SerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGly 460
Db 1321 AGCTCTCTGTGCTGGTGGCTGGAGCGGCTGAGTTGCACGAGCAAAATCATTTGC 1380
Qy 461 PheSerLeuHisTyrGlnLysAlaArgGlyValAspAsnValGluTyrGlnPheAlaVal 480
Db 1381 TTCTCTCTTCACTACCAAAAGGAGGAGTGGACAATGTGGAGTACCAAGTTTGCAGTA 1440
Qy 481 AsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPhe 500
Db 1441 AACAAATGACACACAGAGCTGCAGGTTCGGAGCTTGGAACTTGGAACTTATGAGTTTC 1500
Qy 501 TyrValValAlaTyrSerGlnLeuGlyValAspArgThrSerProAlaLeuValHis 520
Db 1501 TACGTGTGTGCTACTCTCCAGCTGGGGGCGAGCGAATCTCCAGCCAGCGCTTGGTGCAT 1560
Qy 521 ThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSer 540

Db 1561 ACACCTGGACGATGTCTCCAGCGGACGACCCAGCTTACCTTGTTCAGCCCAACCCCTCG 1620
Qy 541 AspIleArgValAlaTrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLys 560
Db 1621 GACATCAGGTTGCAATGGCTGCGCCCTCCAGCTGAGCAATGACAGGTGCTGAAG 1680
Qy 561 TyrLysIleGluTyrGlyLeuGlyLysGluAspGlnValPheSerThrGluValProGly 580
Db 1681 TACAAATAGATAGTACGGTTTGGGGAAGGAAGATCAGGTTTCTCCACCGAGGTGCTCGGA 1740
Qy 581 AsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgIle 600
Db 1741 AATGAGACACACTTACGTTAACTCACTTCAGCCAAACAAAGTGTACCGAGTCCGAGTT 1800
Qy 601 SerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrPro 620
Db 1801 TCAGCTGGCACCTGGCGCTGGCTATGGAGTCCCTTCTCAGTGGATGACGACACACCT 1860
Qy 621 GlyValHisAsnGlnSerHisValProPheAlaProAlaGluLeuLysValArgAlaLys 640
Db 1861 GGTGTGCACACACAGAGCCATGTTCCCTTTGGCCCTTCAGAAATTGAAGGTGAGGGCAAG 1920
Qy 641 MetGluSerLeuValValSerTrpGlnProProHisProThrGlnIleSerGlyTyr 660
Db 1921 ATGGAGTCCCTGTGTGTGTCTATGGCAGCGCCCTCACCCACCCAGATCTCTGTGATAC 1980
Qy 661 LysLeuTyrTrpGlyGluValGlyThrGluGluGluAlaAspGlyAspArgProProGly 680
Db 1981 AAATCTACTGGGAGAGGTGGGAACAGAGGAGGAGGAGATGTTGACCGCCCGCCAGGG 2040
Qy 681 GlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLysLysValLysGln 700
Db 2041 GGTGTGTGAGATCAAGCTTGGGACGTGGCGCCCGTGGAGTGAAGTGAAGTGAAGCAG 2100
Qy 701 TyrGluLeuThrGlnLeuValProGlyArgProTyrGluValLysLeuValAlaPheAsn 720
Db 2101 TATGAATCTGACCCAGTTAGTCCCTGGCAGCGCGTACGAGGTGAAGTCTGTAGCTTTCAAC 2160
Qy 721 LysHisGluAspGlyTyrAlaAlaValTrpLysGlyLysThrGluLysValProThrPro 740
Db 2161 AAACACGAGGAGCGCTACGCTGTGTGTGGAGGGCAGACGAGGAGGGCGCCACGCGCA 2220
Qy 741 AspLeuProIleGlnArgGlyProProLeuProProAlaHisValHisAlaGluSerAsn 760
Db 2221 GACCTGCCTATCCAGAGGGGCGCACCGCTGCTCTGCTCCATGTCACGAGAGTCAAAC 2280
Qy 761 SerSerThrSerIleTrpLeuArgTrpLysProAspPheThrThrValLysIleVal 780
Db 2281 AGCTCCACTTCCATTTGGCTTCGGTGGAAAGAACCCAGACTTTTACCACTGTCAAGATTGTC 2340
Qy 781 AsnTyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTyrTyr 800
Db 2341 AACTACACTGTAGCTTGGCGCCCTGGGGGCTCAGGAATGCTTCCCTGTGTCACTACTAT 2400
Qy 801 ThrSerSerGlyGluAspIleLeuIleGlyLeuLysProPheThrThrLysTyrGluPhe 820
Db 2401 ACCAGCTCTGGAGAACACATTCTCATTTGGCGGCTCGAAACCAATTTACCAAGTACGAGTTT 2460
Qy 821 AlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSer 840
Db 2461 GCGGTACGTCCACCGAGTGGATATGGATGGGCGCTTTTGGCTCCGTCGTAGAACGCTCC 2520
Qy 841 ThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrPro 860
Db 2521 ACCCTGCTGACCGGCTTCAACACCTCTTCTTGACCTGGCGCTGAGCCCTTGACACCA 2580
Qy 861 SerThrValArgLeuHisTrpCysProProThrGluProAsnGlyGluIleValGluTyr 880
Db 2581 TCCACGCTTCGGTTACACTGGTGTCCCGCCACGAGGCCCAATGCTGAGATTGTGGAGTAT 2640
Qy 881 LeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThrThr 900
Db 2641 CTAATTTCTCTACAGCAACAACCAACACCGAGCCCGGAACACCAAGTGGACACTGCTCACCACA 2700

Qy	739	ThrProAspLeuProIleGlnArgGlyProProLeuProProAlaHisValHisAlaGlu	758
Db	2218	GCACCAAGACATGCTATTCAGAGAGGGACCAACCCCTCGCTCCAGCCACGCTCCATGCGAA	2277
Qy	759	SerAsnSerSerThrSerIleTrrpLeuArgTrrpValysProAspPheThrThrVallys	778
Db	2278	TCAACAGAGCTCCACATCATCTGGCTTCGGTGGAAAGCCAGATTTCACCAAGTCAAG	2337
Qy	779	IleValAsnTyrThrValArgPheGlyProTrrpGlyLeuArgAsnAlaSerLeuValThr	798
Db	2338	ATTGTCACTACTGTGGCTTCAGCCCTGGGGGCTCAGGAATGCTTCCCTGGTCAAC	2397
Qy	799	TyrTyrThrSerSerGlyGluAspIleLeuIleGlyLeuLysProPheThrLysTyr	818
Db	2398	TAATTACACCAAGTCTTGAGAAGACATCCTCATTTGGGGCTTGAAGCCATTTCACCAATATC	2457
Qy	819	GluPheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGlu	838
Db	2458	GAGTTTGACGTGAGTCTCAAGCGGTGGACATGATGGGCTTTTCGGCTCTGTGGTGGAG	2517
Qy	839	ArgSerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeu	858
Db	2518	CGCTCCACCTTGCTGACGGGCTTCACACCCCATCCGACCTGGACCTGACGCCCTG	2577
Qy	859	ThrProSerThrValArgLeuHisTrrpCysProProThrGluProAsnGlyGluIleVal	878
Db	2578	ACACCGTCCACGGTTCGGCTGCATGTGTGCCCCCCACACAGAGGCCAACCGGGAGATCGTG	2637
Qy	879	GluTyrLeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTrrpThrLeuLeu	898
Db	2638	GAGTATCTGATCTGTACAGGAGACCAACACGAGGCTTGAGCACAGTGGACTTGGCTC	2697
Qy	899	ThrThrGluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyr	918
Db	2698	ACCACGAGGGAAACATCTTCAGTGTCTGAGTCCATGGCTTGGAGAGCACCTCGGTAC	2757
Qy	919	PhePheLysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAsp	938
Db	2758	TTCTTCAAGATGGGGGCGGCACAGAGGTGGACCTGGGCTTCTCCGCCCTTCACAGGAT	2817
Qy	939	ValIleThrLeuGlnGluThrPheSerAspSerLeuAspValHisAlaValThrGlyIle	958
Db	2818	GTGATCAGCTCCAGGAAGAGCTGTACAGTCTGCTGCACATGACACTCACTGACGGGCATC	2877
Qy	959	IleValGlyValCysLeuLeuGlyLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArg	978
Db	2878	ATCGTGGGTGTCTGCTGGGCTCTCTGCTCTGCTGCTGCATGTGTGCTGGCTGCGC	2937
Qy	979	GlnSerSerHisArgGluAlaLeuProGlyLeuSerSerSerSerGlyThrProGlyAsnPro	998
Db	2938	CGCAGCCCCACAGGGAAATCCCTCCAGGCTGTCTCCACCGCACCCCGGGGAATCCC	2997
Qy	999	AlaLeuTyrThrArgAlaArgLeuGlyProProSerValProAlaHisGluLeuGlu	1018
Db	2998	CGCGTACTACAGCTGGGCTTGGCCCCCCCCAGCCCCCAGCTGCCCATGAATGGAG	3057
Qy	1019	SerLeuValHisProArgProGlnAspTrrpSerProProSerAspValGluAspLys	1038
Db	3058	TCCCTTGTGCACCCCATCCCAAGACTGGTCCCCCGCCACCCCTCAGACGTGGAGGACAGG	3117
Qy	1039	AlaGluValHisSerLeuMetGlyCysValSerAspCysArgGlyHisSerLysArg	1058
Db	3118	GCTGAAGTGCAACGCTTATGGGTGGCGGTGTTTCTGAAGCGCGGAGTCACTCCAAAGA	3177
Qy	1059	LysIleSerTrrpAlaGlnAlaGlyGlyProAsnTrrpAlaGlySerTrrpAlaGlyCysGlu	1078
Db	3178	AAGATCTCTGGGTCAACACAGCGGGCTGAGCTGGGCTGGTCTCTGGGACGGCTGTGAG	3237
Qy	1079	LeuProGlnGlySerGlyProArgProProAlaLeuThrArgAlaLeuLeuProProAlaGly	1098
Db	3238	CTGCCCCAG---GACGCCCCCGCGCTGTGACCGGGCCCTGCTGCCCTCTGCTGCTGGA	3294

Qy	1099	ThrGlyGlnThrLeuLeuGlnAlaLeuValTyrAspGlyIleIysSerAsnGlyArg	1111
Db	3295	ACTGGGCAGACGCTGTGTCTGCAGCGCTCTGGTGTACGACGCCCATAAAGGCGCAATGGGAGG	3354
Qy	1119	LYSLYpSerProAlaCYsArgAAsnGlnValGluAlaGluValIleValHisSerAsp	1138
Db	3355	AAGAAGTCACCCCGCCCTGCAGAAACCAAGTGGAGGCTGAAGTCAATTGTCCACTCTGAC	3414
Qy	1139	PheGlyAlaSerIysGlyCYsProAspLeuHisLeuGlnAspLeuGluProGluGluPro	1158
Db	3415	TTTAGTGCAATCAACGGGAACCTTGACCTCCATCTCCAGACCTGGAGCCTGAGGACCC	3474
Qy	1159	LeuThrAlaGluThrLeuProSerThrSerGlyAlaValAspLeuSerGlnGlyAlaAsp	1178
Db	3475	CTGCCTCCAGAGGCTCTCATCTCATCTCGGGTGTGGGATCCAGGGCAGGGGGCAGCC	3534
Qy	1179	TrpLeuGlyArgGluLeuGlyCYsGlnProThrThrSerGlyProGluArgLeuThr	1198
Db	3535	TGGCTGACAGGAGTGGAGGGTGTGAAGCTGGAGCCCCCGGCCAGACAGACTTACC	3594
Qy	1199	CysLeuProGluAlaAlaSerAlaSerCYsSerCYsSerAspLeuGlnProSerThrAla	1218
Db	3595	TGCTTGCAGAGCGACGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	3654
Qy	1219	IleGluGluAlaProGlyIysSerCYsGlnProProlyAlaLeuCysProLeuThrValSer	1238
Db	3655	CTAGAGGAGACCCCTTGGAGATAGCTGCAGCTCAAAATCCCCCTGCCCTCTAGGAGCAGC	3714
Qy	1239	ProSerLeuProAlaProValSerSerAla 1249	
Db	3715	CCAGGCCTGCCAGATCCCCGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	3747
RESULT 3			
ABK11101			
ID	ABK11101	standard; DNA; 3741 BP.	
XX	AC	ABK111101;	
XX	AC		
XX	AC		
DT	05-JUN-2002	(first entry)	
XX			
DE		DNA encoding human NOV1 protein, homologue of NOPE/PUNC Ig proteins.	
XX			
KW		Human; NOVX-associated disorder; developmental disorder; blood disorder;	
KW		endocrine disorder; vascular disease; gastrointestinal disorder; cancer;	
KW		respiratory disorder; inflammatory disorder; reproductive disorder;	
KW		neurodegenerative disorder; autoimmune disorder; infectious disease;	
KW		cardiovascular disorder; cell signal processing; PUNC immunoglobulin;	
XX		metabolic pathway modulation; NOPE; Ig; NOV1; gene; ds.	
OS		Homo sapiens.	
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..3741	
FT		/*tag= a	
FT		/partial	
FT		/product= "NOV1"	
FT		/note= "This sequence lacks a stop codon"	
XX			
FN	WO200206329-A2.		
XX			
PD	24-JAN-2002.		
XX			
PP	18-JUL-2001; 2001WO-US22709.		
XX			
PR	18-JUL-2000; 2000US-218870P.		
PR	18-JUL-2000; 2000US-218875P.		
PR	18-JUL-2000; 2000US-218901P.		
PR	24-JUL-2000; 2000US-220273P.		
PR	26-JUL-2000; 2000US-220912P.		
PR	27-JUL-2000; 2000US-221233P.		
PR	28-JUL-2000; 2000US-221650P.		
XX			
PA	(CURA-) CURAGEN CORP.		

QY 518 LeuValHisThrLeuAspValProSerAlaAlaProGlnLeuThrLeuSerSerPro 537
Db 1555 CTGGTGCAACACTGGATGATCTCCAGTCAGCAGCAGCCAGCTCTCCCTGTCAGGCC 1614
QY 538 AenProSerAspIleArgValAlaTrpLeuProLeuProSerSerLeuSerAenGlyGln 557
Db 1615 AACCTTCGACATCAGGGTGGCGTGGCTGCCCTGCCCTGCCCTGCCCTGAGCAATGGGCG 1674
QY 558 ValLeuLysTrpLysIleGluTrpGlyLeuGlyLysGluAspGlnValPheSerThrGlu 577
Db 1675 GTGGTGAAGTACAAGATAGAAACGGTTGGGAAGAGAGATCAGATTTCCTACTGAG 1734
QY 578 ValProGlyAsnGluThrGlnLeuThrLeuAenSerLeuGlnProAenLysValTrpArg 597
Db 1735 GTGCGAGGAAATCAGACACAGCTTATGCTGAACCTGCTTTCAGCCAAACAGGTGTATCGA 1794
QY 598 ValArgIleSerAlaGlyThrGlyAlaGlyTrpGlyValProSerGlnTrpMetGlnHis 617
Db 1795 GTACGGATTTCGGCTGGTACAGACCGCGCTTTCGGGGCCCTCCAGTGGATGATCAC 1854
QY 618 ArgThrProGlyValHisAenGlnSerHisValProPheAlaProAlaGluLeuLysVal 637
Db 1855 AGGAGCCCGAGTATGCACACACAGGCCATGTCCTTTTGGCCCTGCAGATTGAGGTG 1914
QY 638 ArgAlaLysMetGluSerLeuValValSerTrpGlnProProProHisProThrGlnIle 657
Db 1915 CAGGCAAAAGATGGAGTCCCTGCTGCTGTCATGGCAGCCACCCCTCACCCACCCAGATC 1974
QY 658 SerGlyTrpLysLeuTrpGlyGluValGlyThrGluGluGluAlaAspGlyAspArg 677
Db 1975 TCTGGCTACAAACTATATTGGCGGAGGTGGGGCTGAGGAGGAGGCCAATGGCGATCGC 2034
QY 678 ProProGlyGlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLysLysLys 697
Db 2035 CTGCCAGGGGGCGTGGACACACAGGCTTGGGATGTGGGGCTGTCCGGCTCAAGAGAAA 2094
QY 698 ValLysGlnTrpGluLeuThrGlnLeuValProGlyArgTrpTrpGlyValLysLeuVal 717
Db 2095 GTGAAGCATATGAGCTGACCCAGCTAGTCCCTGGCGGCTGTGTCGAGGTGAAGCTGGT 2154
QY 718 AlaPheAsnLysHisGluAspGlyTrpAlaAlaValTrpLysGlyThrGluLysAla 737
Db 2155 GCTTTTCACAAACATGAGATGGCTATGACAGCAGTGTGGAGGGCAAGACGAGAGAGCGC 2214
QY 738 ProThrProAspLeuProIleGlnArgGlyProProLeuProProAlaHisValHisAla 757
Db 2215 CCGGCACCAAGACATGCTTATCCAGAGGGGACACCCCTGCCCTCCAGCCACCTCCATGGC 2274
QY 758 GluSerAsnSerSerThrSerIleTrpLeuArgTrpLysLeuProAspPheThrThrVal 777
Db 2275 GAATCAAAACAGCTCCACATCCATCTGGCTTCGGTGGAAAAAGCCAGATTTCACACAGTC 2334
QY 778 LysIleValAsnTrpThrValArgPheGlyProTrpTrpGlyLeuArgAenAlaSerLeuVal 797
Db 2335 AAGATTGTCAACTACATCTGTGGCTTTCAGCCCTTCGGGGCTCAGGAAATGCCCTCCCTGGTC 2394
QY 798 ThrTrpTrpThrSerSerGlyLysAspIleLeuIleGlyLysLeuLysProPheThrLys 817
Db 2395 ACCATATAC---AGTTCTGAGAAACATCTCTATTTGGCGGCTTGAAGCCATTCACAAA 2451
QY 818 TyrGluPheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValVal 837
Db 2452 TAGCAGTTTTCAGTGCAGTCTCAGCGGTGGACATGATGGGCTTTCGGCTCTGTGGTG 2511
QY 838 GluArgSerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerPro 857
Db 2512 GAGCGCTCCACCCCTGCTGACCGTCCCTCCACACCCCTCCAGACCTCGGACTGAGCCCT 2571
QY 858 LeuThrProSerThrValArgLeuHisTrpCysProProThrGluProAenGlyLysIle 877
Db 2572 CTGACACCGTCCACCGTGTGGCTGCACTGGTGGCTGGTGGCTGGTGGCTGGTGGCTGGT 2631
QY 878 ValGluTrpLysLeuIleLeuTrpSerAsnAenHisThrGlnProGluHisGlnTrpThrLeu 897

Db 2632 GTGGAGTATCTGATCTCTGTATCAGGCAACACACAGCAGCTCTGAGCACCTGAGCTG 2691
QY 898 LeuThrThr-----GluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAsp 915
Db 2692 CTCACCAAGAGGTGAGGGAAACATCTTCACTGCTGAGGTCCATGGCTGAGAGGCAC 2751
QY 916 ThrArgTrpPhePheLysMetGlyAlaArgThrGluValGlyProGlyProPheSerArg 935
Db 2752 ACTCGTACTTCTTCAAGATGGGGCGGCACAGAGGTGGGACCTGGGCTTCTCCGCGC 2811
QY 936 LeuGlnAspValIleThrLeuGlnThrPheSerAspSerLeuAspValHisAlaVal 955
Db 2812 CTGAGGATGTGATCAGCTCCAGGAGAAAGTGTGACAGCTCGCTGGACATCCTCAGTC 2871
QY 956 ThrGlyIleIleValGlyValCysLeuGlyLeuLeuCysLeuAlaCysMetCysAla 975
Db 2872 ACGGGCATCATCTGGGTGTCTGCTGGGCTCTCTGCTCTCTGGCTTGCATGTGTGCT 2931
QY 976 GlyLeuArgGlnSerSerHisArgGluAlaLeuProGlyLeuSerSerSerGlyThrPro 995
Db 2932 GGCCTGGCGCGCAGCCCCACAGGGAAATCCCTCCAGGCTGTCTCCACCGCCACCC 2991
QY 996 GlyAenProAlaLeuTrpThrArgAlaArgLeuGlyProProSerValProAlaAlaHis 1015
Db 2992 GGGAAATCCCGGCTGTACTCCAGAGCTCGGCTGGCCCTCCCGCCAGCCCTCGCTGCCAT 3051
QY 1016 GluLeuGluSerLeuValHisProArgProGlnAspTrpSerProProSerAspVal 1035
Db 3052 GAATTGGAGTCCCTTGTGCACCCCATCCAGAGACTGTGCTCCCGCCACCTCTCAGACGTG 3111
QY 1036 GluAspLysAlaGluValHisSerLeuMetGlyGlySerValSerAspCysArgGlyHis 1055
Db 3112 GAGGACAGGGCTGAAGTGCACAGCTTATGGTGGCGGTGTTCTGAAGGCGGAGTCAAC 3171
QY 1056 SerLysArgLysIleSerTrpAlaGlnAlaGlyGlyProAenTrpAlaGlySerTrpAla 1075
Db 3172 TCCAAAAGAAAGTAAAGT---GCTCAACCAAGCGGCTGAGCTGGGCTGTGTTCTCGGCA 3228
QY 1076 GlyCysGluLeuProGlnGlySerGlyProArgProAlaLeuThrArgAlaLeuLeuPro 1095
Db 3229 GCTGTGAGCTGCCCCAG---GCGGGCCCGCGCGGCTCTGACCCGGGCTCTGCTGCC 3285
QY 1096 ProAlaGlyThrGlyGlnThrLeuLeuLeuGlnAlaLeuValTyrAspGlyIleLysSer 1115
Db 3286 CCTGTGTAACCTGGGACAGCTGTGTGCTGCAGGTTCCTCTGCTCTGAT-----CAGGCG 3339
QY 1116 AenGlyArgLysLysProSerProAlaCysArgAenGlnValGluAlaGluValIleVal 1135
Db 3340 AATGGGAGAAAGATCACCCCGCTGCAGGAACCAAGGTGGAGGCTGAAGTCAATTGTC 3399
QY 1136 HisSerAspPheGlyAlaSerLysGlyCysProAspLeuHisLeuGlnAspLeuGluPro 1155
Db 3400 CACTCTGACTTTAGTGCATCTAACCGGAACCTGACCTCCATCTCCAGACCTGAGGCT 3459
QY 1156 GluGluProLeuThrAlaGluThrLeuProSerThrSerGlyValAlaValAspLeuSerGln 1175
Db 3460 GAGGACCCCTGCCCTCCAGAGGCTCTCATCTCATCTCGGGTGTGGGATCCAGGGCAG 3519
QY 1176 GlyAlaAspTrpLeuGlyArgGluLeuGlyGlyCysGlnProThrThrSerGlyProGlu 1195
Db 3520 GGGGCGAGCTGCTGGACAGAGGAGTGGGAGGCTGTGAGCTGGCAGCCCGCCGGGCGCAGAC 3579
QY 1196 ArgLeuThrCysLeuProGluAlaAlaSerAlaSerCysSerCysSerAspLeuGlnPro 1215
Db 3580 AGACTTACTGCTGCCAGAGCAGCCAGTGTCTCTGCTCTCTACCCGACCTCCAGGCCA 3639
QY 1216 SerThrAlaIleGluGluAlaProGlyLysSerCysGlnProLysAlaLeuCysProLeu 1235
Db 3640 GCGGAGGTGCTAGAGAGAGACCCCTGGAGATAGTCTCCAGCTCAATATCCCTTCCCTCTA 3699
QY 1236 ThrValSerProSerLeuProArgAlaProValSerSerAla 1249

Db	961	GCGCTCTCGCGGACGCGGGCCAGACCGCGCGCTTCGTGTGCCGGGGTCCGGGGAGCCA	1020
Qy	362	ArgProAlaLeuHieTrpLeuHieAspGlyIleProLeuArgProAsnGlyArgValIys	381
Db	1021	CGGCCCGCGCTGCATCGCTGCACGCGGATCCCGTTGCGACCAATGGCGCGGTCAAG	1080
Qy	382	ValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyr	401
Db	1081	GTGAGGCGCGTGGCGGAGCTTGGTCACTCAGATCGGCTGCAGGACCGTGGCTAC	1140
Qy	402	TyrGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaProLeuAlaVal	421
Db	1141	TACCAGTGGGTAGCAGAAAACAGCGCGGAACTGCTGTGCGCTGCGCCCTGCGGTA	1200
Qy	422	ValValArgGluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSer	441
Db	1201	GTGGTGGCGAGGGGCTGCCAGGCGCCCGACTCGGGGTCAACGACCGCGCTGAGCAGC	1260
Qy	442	SerSerValLeuValAlaTrpGluArgProGluLeuHieSerGluGlnIleGlyPhe	461
Db	1261	TCTCTGTGCTGGTGGCTGGAGCGGCTGAGTTGCACGCGAGCAATCATTTGGCTTC	1320
Qy	462	SerLeuHieTrpGlnLysAlaArgGlyValAspAsnValGluTrpGlnPheAlaValAsn	481
Db	1321	TCTCTTCACTACCAAAAGGCAAGGGAGTGGCAATGTGGAGTACCAGTTTGCAGTAAC	1380
Qy	482	AsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPheTyr	501
Db	1381	AATGACACACACAGAGTCCAGGTTCCGGACCTCGGAACCCCAACACGGAATTATGATTCTAC	1440
Qy	502	ValValAlaTrpSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHieThr	521
Db	1441	GTGGTGGCTTACTTCCAGCTGGGGGCGAGCGCAACTCCAGCCAGCCCTGGTGATACA	1500
Qy	522	LeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAsp	541
Db	1501	CTGGACGATGTCCCGAGCGCAGCACCCAGCTTACCTTGTCCAGCCCAACCCCTCGGAC	1560
Qy	542	IleArgValAlaTrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLysTyr	561
Db	1561	ATCAGGGTGGCATGCTGCCCTGCCCTCCAGCTTGAGCAATGGACAGGTGCTGAAGTAC	1620
Qy	562	LysIleGluTrpGlyLeuGlyGlyGluAspGlnValPheSerThrGluValProGlyAsn	581
Db	1621	AAGATAGAGTACGGTTTGGGGAAGGAAGATCAGGTTTTCTCCAGCGGTGCTGGAAT	1680
Qy	582	GluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgIleSer	601
Db	1681	GAGACACAACTTACGTTAACTCACTTCAGCCAAACAAAGTGTACCGAGTCCGGATTCA	1740
Qy	602	AlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGlnHieArgThrProGly	621
Db	1741	GCTGGCACTGGCGCTGGCTATGGAGTCCCTTCTCAGTGGATGTCAGCACAGGACACTGGT	1800
Qy	622	ValHieAsnGlnSerHieValProPheAlaProAlaGluLeuLysValArgAlaLysMet	641
Db	1801	GTGCAAAACAGAGACCAATGTTCCCTTTCCTGCGGCTGAGAAATGAAGGTGAGGCAAGATG	1860
Qy	642	GluSerLeuValValSerTrpGlnProProHisProThrGlnIleSerGlyTyrLys	661
Db	1861	GAGTCCCTGGTGGTGTATGGAGCGGCCCTCTACCCCAACCCAGATTTCTGGATACAA	1920
Qy	662	LeuTyrTrpGlyGluValGlyThrGluGluAlaAspGlyAspArgProProGlyGly	681
Db	1921	CTCTACTGGGAGAGGTGGAAACAGAGGAGGAGGAGATGGTGACCGCCCCCAGGGGGT	1980
Qy	682	ArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLysLysValLysGlnTyr	701
Db	1981	CSTGGAGATCAAGCTTGGACGTCCGGGCGCGTGGCGGCTGAAGAGAAAGTGAAGCAGTAT	2040
Qy	702	GluLeuThrGlnLeuValProGlyArgProTyrGluValLysLeuValAlaPheAsnLys	721
Db	2041	GAACCTGACCAGTTAGTTCCTGGCAGGCGGTACGAGGTGAAGCTCGTAGCTTTTCAACAA	2100
Qy	722	HieGluAspGlyTyrAlaAlaValTrpLysGlyLeuThrGluLysValProThrProAsp	741
Db	2101	CAGAGACCGCTACGCTGCTGTGTGAAGGGCAGACGGAAGAGCGCCCAACGCGCAGAC	2160
Qy	742	LeuProIleGlnArgGlyProProLeuProProAlaHieValHieAlaGluSerAsnSer	761
Db	2161	CTGCTTATCCAGAGGGGGCCACCGCTGCTTCTGCTCCATGTCACGAGAGTCAACAGC	2220
Qy	762	SerThrSerIleTrpLeuArgTrpLysLysProAspPheThrThrValLysIleValAsn	781
Db	2221	TCCACTTCCATTTGGCTTCGGTGGAAAGAGCCAGACTTTTACCACTGTCAAGATTGTCAAC	2280
Qy	782	TyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTyrTyrThr	801
Db	2281	TACACTGTACGCTTCGCGCCCTCGGGCTCAGGAATGCTTCCCTGTGTCACTACTATACC	2340
Qy	802	SerSerGlyGluAspIleLeuIleGlyLeuLysProPheThrTrpLysTyrGluPheAla	821
Db	2341	AGCTCTGGAGAGACATTTCTCATTTGGCGGCTGAAACCAATTTACCAAGTACGAGTTTGGC	2400
Qy	822	ValGlnSerHieGlyValAspMetAspGlyProPheGlySerValValGluAspSerThr	841
Db	2401	GTACAGTCCCAACGAGTGGATATGGATGGGCGCTTTGGCTCCGTCTGTAGAACGCTCCACC	2460
Qy	842	LeuProAspAspProSerThrProProSerAspLeuArgLeuSerProLeuThrProSer	861
Db	2461	CTGCTGACCGGCTTCCACACTCTCTTGACCTGCGCTGAGCCCTCCGACACCATCC	2520
Qy	862	ThrValArgLeuHieTrpCysProProThrGluProAsnGlyGluIleValGluTyrLeu	881
Db	2521	ACCGTTCCGTTACACTGGTGTCCCCCACGAGGCCAATGGTGAGATTGTGGAGTACTA	2580
Qy	882	IleLeuTyrSerAsnAsnHieThrGlnProGluHieGlnTrpThrLeuLeuThrThrGlu	901
Db	2581	ATTCTCTACAGCAACACACACACCGCCGGAACACCACTGCTGCACTGCTCACCACAG	2640
Qy	902	GlyAsnIlePheSerAlaGluValHieGlyLeuLysSerAspThrArgTyrPhePheLys	921
Db	2641	GGAAACATCTTCAGTGCAGAGGTCCATGGGCTTAGAGAGTGACACTCGGTATTTCTTCAAG	2700
Qy	922	MetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIleThr	941
Db	2701	ATGGAGCCCGCACAGAGGTGGGCGCTTGGGCGCTTTTCCCGCTTGCAGGATGTGATTACT	2760
Qy	942	LeuGlnGluThrPheSerAspSerLeuAspValHis	953
Db	2761	CTGCAAGAGACATTTCTCAGACTCTCTGGATGTGCAC	2796
RESULT 5			
Db	ABZ11219		
ID	ABZ11219	standard; cdna; 3450 BP.	
XX	XX	ABZ11219;	
AC	ABZ11219;		
XX	XX	20-JAN-2003 (first entry)	
DT	XX	Human polynucleotide SEQ ID NO 101.	
DE	XX		
XX	XX	Human; genome mapping; gene therapy; food supplement; virus; fungus;	
KW	KW	cell-proliferative disorder; neurodegenerative disease; bacterial;	
KW	KW	Parkinson's disease; Alzheimer's disease; autoimmune disease;	
KW	KW	multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;	
KW	KW	arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;	
KW	KW	antiparkinsonian; antidiabetic; immunosuppressive; dermatological;	
KW	KW	haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;	
XX	XX	antiarthritic; gene; ss.	
OS	XX	Homo sapiens.	
XX	XX	WO200270539-A2.	
PN	XX		
XX	XX		

PD 12-SEP-2002.
 XX 05-MAR-2002; 2002WO-US05095.
 XX 05-MAR-2001; 2001US-0799451.
 XX (HYSB-) HYSEQ INC.
 XX Tang YT, Zhou P, Goodrich RW, Auendi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
 PI Wehrman T, Wang J, Wang D, Dmanac RT;
 XX WPI; 2002-759812/82.
 DR P-PSDB; ABP69002.
 XX
 PT New polynucleotides comprising sequences assembled from expressed
 PT sequence tags (ESTs), useful for treating cell-proliferative,
 PT neurodegenerative, autoimmune genetic, myeloid or lymphoid, or
 PT platelet or coagulation disorders -
 XX
 PS Claim 1; SEQ ID NO 101; 1012pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated polynucleotide (I) comprising a
 CC nucleotide sequence selected from any of 948 sequences
 CC (ABZ1119-ABZ12066) or their mature protein coding portion, active domain
 CC coding protein or complementary sequences. The polynucleotides are useful
 CC for identifying expressed genes or for physical mapping of human genome.
 CC The encoded polypeptides (ABP68902-ABP69849) are useful as molecular
 CC weight markers, as a food supplement, for generating antibodies, in
 CC medical imaging, screening and diagnostic assays and for treating
 CC cell-proliferative disorders (cancer), neurodegenerative diseases
 CC (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple
 CC sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid
 CC disorders, platelet or coagulation disorders, wound, burns, incision,
 CC ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,
 CC parasitic), arthritis, etc.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 3450 BP; 681 A; 1050 C; 1067 G; 652 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,46e-200 Length: 3450
 Score: 4361.50 Matches: 822
 Percent Similarity: 92.55% Conservative: 35
 Best Local Similarity: 88.77% Mismatches: 68
 Query Match: 66.04% Indels: 1
 DB: 24 Gaps: 1
 US-09-754-997A-2 (1-1252) x ABZ11219 (1-3450)
 QY 324 AlaAlaAlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeu 343
 DB 163 TCAGCGTGTGCGCCCTCGCGCCCTTAGCGGCTCCCGCCATCATCAGCGCGCGCGAGGCGCTG 222
 QY 344 SerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgPro 363
 DB 223 TCSCGGAACGGCGGAGCAGACGCGCGCTTCGTGTGCGCGCTGCGGAGCGCGCGCCCA 282
 QY 364 AlaLeuHisTrpLeuHisAspGlyIleProLeuArgProAlaGlyArgValValGln 383
 DB 283 GCGCTGCGCTGCTGCACACACGGCGCGCGCTGCGGCCCAACGCGCGCTCAAGGTCCTAG 342
 QY 384 GlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrTyrGln 403
 DB 343 GCGCGCGGTGCGAGCTGTGTCATCACACATCGGCTGCAGGACGCGCGCTACTACACAG 402
 QY 404 CysValAlaGluHisSerAlaGlyThrAlaCysAlaAlaAlaProLeuAlaValVal 423
 DB 403 TCGTGCTGCTGAGAACACGCGGGGAATGCGCTGCGCTGCGCGCTGCGCGCGGTGGTG 462
 QY 424 ArgGluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSer 443

DB 463 CCGCAGGGGCTGCCCCAGCGCCCCCGGGGTCATCTGCTACGCCACTGAGCAGCTCCGCT 522
 QY 444 ValLeuValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleIleGlyPheSerLeu 463
 DB 523 GTGTGTGGTGGCTGGAGCGGCGCCGAGATGCACAGCGAGCAGATCATCGGCTTCTCTCTC 582
 QY 464 HisTyrGlnLysAlaArgGlyValAspAsnValGluTyrGlnPheAlaValAsnAsnAsp 483
 DB 583 CACTACCAAGAGCGCGGCGCATGGACAATGTGGAAATACAGTTTGCAGTGAACACGAC 642
 QY 484 ThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPheTyrValVal 503
 DB 643 ACCCAGAACTACAGGTCGGGACCTGGAAACCCCAACACAGATATAGAGTTCTACGTGG 702
 QY 504 AlaTyrSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThrLeuAsp 523
 DB 703 GCTTACTCCAGCTGGAGCGCAGCGCACCTCCACCCAGCACTGGTGCACACACTGGAT 762
 QY 524 AspValProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAspIleArg 543
 DB 763 GATGTCCCGAGTGCAGCACCCCGAGCTCTCCCTGTCCAGCCCAACCCCTTCGGACATCAG 822
 QY 544 ValAlaTrpLeuProLeuProSerSerSerSerSerSerSerSerSerSerSerSerSer 563
 DB 823 GTGGCGTGGCTGCTCCCTGCCCCCGCCAGCAATGGCGAGTGGTGAAGTACAGATA 882
 QY 564 GluTyrGlyLeuGlyLysGluAspGlnValPheSerThrGluValProGlyAsnGluThr 583
 DB 883 GAATACGGTTTGGAAAGGAGATCAGATTCTCTACTAGGTGGGAGGAAATGAGACA 942
 QY 584 GlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgIleSerAlaGly 603
 DB 943 CAGCTTATGTCTGAACTCGCTTCAGCCAAACAAAGGTATCGAGTACGGATTCGGCTGG 1002
 QY 604 ThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrProGlyValHis 623
 DB 1003 ACAGCAGCGCGCTTCGGGGCCCCCTCCCAAGTGGATGCATCAGAGGACGCCAGTATGCA 1062
 QY 624 AsnGlnSerHisValProPheAlaProAlaGluLeuLysValArgAlaLysMetGluSer 643
 DB 1063 AACCCAGGCCATGTCTCTTTTCCCTTCGCGAGTTGAGGTGCGAGCAAGATGGAGTCC 1122
 QY 644 LeuValValSerTrpGlnProProHisProThrGlnIleSerGlyTyrLysLeuTyr 663
 DB 1123 CTGGTGTGTGTGCGAGCCACCCCTCACCACCCACCCAGATCTCTGGCTACAAATATAT 1182
 QY 664 TrpGlyGluValGlyThrGluGluAlaAspGlyAspArgProProGlyGlyArgGly 683
 DB 1183 TGGCGGGAGGTGGGGGCTGAGGAGGAGGCCAATGGCGATCGCTGCGAGGGGCGCGTGA 1242
 QY 684 AspGlnAlaTrpAspValGlyProValArgLeuLysValLysValLysGlnTyrGluLeu 703
 DB 1243 GACCCAGGCTTGGAGTGTGGGGCTGTCCGGCTCAGAGAGAGTGAAGCAGTATAGCTG 1302
 QY 704 ThrGlnLeuValProGlyArgProTyrGluValLysLeuValAlaPheAsnLysHisGlu 723
 DB 1303 ACCCAGCTAGTCCCTGGCGCGCTGTACGAGGTGAGCTGTGTGCTTTCACCAACATGAG 1362
 QY 724 AspGlyTyrAlaAlaValTrpLysGlyLysThrGluLysAlaProThrProLeuPro 743
 DB 1363 GATGCTATGCAGCATGTGGAAAGGCGCAAGACGAGAGAGCGCGCGCACGACATGCT 1422
 QY 744 IleGlnArgGlyProProLeuProProAlaHisValHisAlaGluSerSerSerSer 763
 DB 1423 ATCCAGAGGGAGCACCCCTGCTCCAGCCACGCTCAGTGGGAATCAACAGCTCCACA 1482
 QY 764 SerIleTrpLeuArgTrpLysLysProAspPheThrThrValLysIleValAsnTyrThr 783
 DB 1483 TCCATCTGCTTGGTGGGAAAAGCCAGATTTCCACAGTCAAGATTGTCAACTACTACT 1542
 QY 784 ValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTyrTyrThrSerSer 803

Db 1543 GTGCGCTTTCAGCCCTGGGGGCTCAGGAATGCTCCCTGGTCACTATTACACCAAGTTCT 1602
Qy 804 GlyGluAapIleLeuIleGlyGlyLeuIleValPheThrLeuValGluPheAlaValGln 823
Db 1603 GGAGAAGACATCTCTATTTGGGGCTTGAAGCCATTTCAACAAATAGAGTTTGGAGTGAG 1662
Qy 824 SerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSerThrLeuPro 843
Db 1663 TCTCAGCGGTGGACATGATGGTGGCTTTTGGCTCTGTGGTGGAGGGCTCCACCTGGCT 1722
Qy 844 AspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrProSerThrVal 863
Db 1723 GACCGGCTTCACACCCCTCCGACCTGCGACTGAGCCCTGACACCGCTCCACGGT 1782
Qy 864 ArgLeuHisTrpCysProProThrGluProAsnGlyGluIleValGluTyrrLeuIleLeu 883
Db 1783 CGGCTGCACTGGTGGCTCCCAAGAGCCCAAGGGAGATCGTGAGTATCTGATCTG 1842
Qy 884 TyrSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThrThrGluGlyAsn 903
Db 1843 TACAGCAGCAACACACAGCAGCTGAGCACACAGTGGACCTTCTCACCAGCGAGGAAAC 1902
Qy 904 IlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrrPhePheIleValMetGly 923
Db 1903 ATCTTCAGTGTGAGTCCATGGCTTGGAGAGCGACACTCGGTACTTCTTCAAGATGGG 1962
Qy 924 AlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIleThrLeuGln 943
Db 1963 GCGCGACAGAGGTGGACCTGGGCTTCTCCCGCTGACAGATGTGATCAGCTCCAG 2022
Qy 944 GluThrPheSerAspSerLeuAspValHisAlaValThrGlyIleIleValGlyValCys 963
Db 2023 GAGAAGCTCTCAGACTCGCTGGACATGCACTCAGTCACGGGATCATCTGGGTGTCTGC 2082
Qy 964 LeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGlnSerSerHisArg 983
Db 2083 CTGGGCTCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2142
Qy 984 GluAlaLeuProGlyLeuSerSerGlyThrProGlyAsnProAlaLeuThrThrArg 1003
Db 2143 GAATCCCTCCAGGCTGTCTCCACCCGCGCAATCCCGGGAATCCCGGCTGTACTCCAG 2202
Qy 1004 AlaArgLeuGlyProProSerValProAlaAlaHisGluLeuGluSerLeuValHisPro 1023
Db 2203 GCTGGCTTGGCCCCCCCCAGCCCCCAGCTGCCCATGANTGAGTCCCTTGTGACCCC 2262
Qy 1024 ArgProGlnAspTrpSerProProSerAspValGluAspLeuAlaGluValHisSer 1043
Db 2263 CATCCCCAGGACTGGTCCCGCCACCTCAGACGCTGAGGACAGGCTGAAGTGACACAGC 2322
Qy 1044 LeuMetGlyGlySerValSerAspCysArgGlyHisSerIleArgIleSerTrpAla 1063
Db 2323 CTATGGGTGGGGGTGTTCTGAAGCCCGAGTCACTCCAAAGAAAGATCTCTGGGGCT 2382
Qy 1064 GlnAlaGlyGlyProAsnTrpAlaGlySerTrpAlaGlyCysGluLeuProGlnGlySer 1083
Db 2383 CAACCAAGCGGCTGAGCTGGGCTGTTCTGGGAGGCTGTGAGTGGCCCCAG--GCA 2439
Qy 1084 GlyProArgProAlaLeuThrArgAlaLeuLeuProProAlaGlyThrGlyGlnThrLeu 1103
Db 2440 GGGCCCCGGCGCTCTGACCCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2499
Qy 1104 LeuLeuGlnAlaLeuValTrpAspGlyIleIleValSerAsnGlyArgIleValProSerPro 1123
Db 2500 TTGCTGACGGCTCTGTGTGACGACGCAATAAGGGCAATGGGAGGAAGATCACCCTCCA 2559
Qy 1124 AlaCysArgAsnGlnValGluAlaGluValIleValHisSerAspPheGlyAlaSerIle 1143
Db 2560 GCTGCGAGAACAGGCTGAGGCTGAGTCATGTTGCTCACTCTGACTTTAGTGCATTAAC 2619
Qy 1144 GlyCysProAspLeuHisLeuGlnAspLeuGluProGluGluProLeuThrAlaGluThr 1163
Db 2620 GGGAAACCTGACTCTCATCTCCAAAGACCTGGAGCCTGAGGACCCCTGCTCCAGAGGCT 2679

RESULT 6

AAI67203

ID AAI67203 standard; DNA; 2223 BP.

XX

AC AAI67203;

XX

DT 11-FEB-2002 (first entry)

XX

DE Nucleotide sequence of GSK gene id 27142.

KW

KW Peptide hormone; antidiabetic; anorectic; antianorectic; antiasthmatic;

KW antidepressant; nootropic; neuroprotectant; hypotensive; hypertensive;

KW cytosolic; cerebroprotective; vasotropic; human; ds.

XX

OS Homo sapiens.

XX

FN WO200172961-A2.

XX

PD 04-OCT-2001.

XX

PF 22-MAR-2001; 2001WO-US09226.

XX

PR 24-MAR-2000; 2000US-192158P.

PR

PR 28-MAR-2000; 2000US-192668P.

PR

PR 27-APR-2000; 2000US-200166P.

XX

PA (SMIK) SMITHKLINE BEECHAM CORP.

PA

(SMIK) SMITHKLINE BEECHAM PLC.

XX

PI Agarwal P, Murdoch PR, Rizvi SK, Smith RP, Xiang Z, Kabnick KS;

PI Lai Y;

XX

WPI; 2001-639223/73.

DR

P-PSDB; AAG65913.

XX

PT Isolated polypeptides, which may be peptide hormones, which are

PT

PT identified by high throughput genome-based biology which identifies

PT

PT genes and gene products as therapeutic targets for treatment of

PT

PT diseases such as diabetes and cancer

XX

PS Claim 2; Page 54-55; 99pp; English.

XX

CC The invention provides polypeptides (AAG65886-65918) which may be peptide

CC

CC hormones (including insulin, growth hormones, chemokines, cytokines,

CC

CC neuropeptides, integrins, kallikreins, laminins, melanins, natriuretic

CC

CC hormones, neuropepsin, pituitary hormones, pleiotrophins, prostaglandins,

CC

CC secretogranins, selectins, thromboglobulins, thymosins) identified by

CC

CC high throughput genome-based biology and polynucleotides (AAI67176-67208)

CC

CC encoding them. The polypeptides can be expressed by standard recombinant

CC

CC methodology. The polypeptides are useful in the treatment of disease such

CC

CC as diabetes, breast-, prostate-, colon cancer and other malignant tumors,

CC

CC hyper- and hypotension, obesity, bulimia, anorexia, growth abnormalities,

CC

CC asthma, manic depression, dementia, delirium, mental retardation,
 CC Huntington's disease, Tourette's syndrome, schizophrenia, growth, mental
 CC or sexual development disorders, and dysfunctions of the blood cascade
 CC system including those leading to stroke. The polynucleotides may be used
 CC as diagnostic reagents through detecting mutations in the associated gene
 CC and for chromosome localization and for tissue expression studies. The
 CC polypeptides and polynucleotides may also be used as vaccines.

XX SQ Sequence 2223 BP; 440 A; 660 C; 735 G; 388 T; 0 other;

Alignment Scores:

Pred. No.: 2,84e-143 Length: 2223
 Score: 3167.50 Matches: 620
 Percent Similarity: 90.33% Conservative: 34
 Best Local Similarity: 85.64% Mismatches: 54
 Query Match: 47.96% Indels: 18
 DB: 22 Gaps: 2

US-09-754-997A-2 (1-1252) x AAI677203 (1-2223)

Qy 17 LeuLeuSerAlaArgGlyGluLeuProGlnGluThrThrValLysLeuSerCys 36
 Db 61 CTCTCTCTATGCAGGGAGCTGCTGTGCCCCCAGGAGAGCTGTGGAGCTGAGCTGT 120
 Qy 37 AspGluGlyProLeuGlnValLysLeuGlyProGlnGlnAlaValValLeuAspCysThr 56
 Db 121 GGAGTGGGGCCACTGCAGGTATCTCGGGCCAGAGCAGCTGCAGTGTCTAAACTGTAGC 180
 Qy 57 LeuGlyAlaThrAlaAlaGlyProProThrArgValThrTrpSerLysAspGlyAspThr 76
 Db 181 CTGGGGCTCTGCTGCGCTGAGCCCCCAGGAGGAGCTGTGGAGCAAGGATGGGACACC 240
 Qy 77 ValLeuGluHisGluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerPro 96
 Db 241 CTGCTGAGGACAGCACCTTACCTCTGCTGCCCATGTTCCCTGTGGCTGTGCCAGCCA 300
 Qy 97 LeuGlnGlnGluAspSerAspGluAlaLeuArgLysValThrGluGly 116
 Db 301 CTAGCACCCCAATGGCAGTACGAGT--CAGTCCCTGAGGCTGTGGGG-GTCATTGAAGGC 357
 Qy 117 SerTyrSerCysLeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaVal 136
 Db 358 AACTATTTCGTGCTAGCCACCGCCCCCTGGAGTGTCTGGCCAGCCAGACTGTCTGCTC 417
 Qy 137 LysLeuAlaThrLeuGluAspPheSerLeuHisProGlnSerGlnLysValGluAsn 156
 Db 418 AAGCTTCCACACTCGCAGACTTCTCTGACCCCGGAGTCTCAGCGTGGAGGAGAAC 477
 Qy 157 GlyThrAlaArgPheGluCysHisThrLysGlyLeuProAlaProLysLeuThrTrpGlu 176
 Db 478 GGGACAGCTCGCTTGTGAGTGCCACATTGAAGGGCTGCAGCTCCCATCTACTTGGGAG 537
 Qy 177 LysAspGlnValThrValProGluGluProArgLysLeuThrLeuProLysTrpLeuLeu 196
 Db 538 AAGACCCAGGTGACATTGCTCGAGAGCTCGGCTCATCTGCTTCCCAACGGCGCTT 597
 Qy 197 GlnLysLeuAspValGlnAspSerAspAlaGlySerTyrArgCysValAlaThrAsnSer 216
 Db 598 CAGATCTGTGATGTCAGGAGAGTATGCAAGGCCCTTACCCTGCTGGGCCACCACTCA 657
 Qy 217 AlaArgGlnArgPheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGlu 236
 Db 658 GCTCGCCAGCACTTACGACAGGAGGCTTACTAGTGTGGCCACAGAGGCTCCCTGGCG 717
 Qy 237 AlaThrArgGlyGlnAspValValLysLeuAlaAlaProGlnAsnThrThrValValSer 256
 Db 718 TCCACACAGGGGGCAGGAGCTGCTGCTTGTGGCAGCCCCCAGAGAACACCAAGTGTCT 777
 Qy 257 GlyGlnAsnValValMetGluCysValAlaSerAlaAspProThrProPheValSerTrp 276
 Db 778 GGCCAGAGTGTGGTATGATGTGGCTCAGCTGACCCACCCCTTTTGTGTCTCTGG 837
 Qy 277 ValArgGlnAspGlyLysProLysSerThrAspValLysValLeuGlyArgThrAsnLeu 296

Db 838 GTCCGCAAGAGCGGGAAGCCCATCTCCACAGATGTATCTGTCTGGGGCGCAACACCTA 897
 Qy 297 LeuLeuAlaSerAlaGlnProArgHisSerGlyValTyrValCysArgAlaAsnLysPro 316
 Db 898 CTAAATTGCCAACCGGAGCGCTGGCACTCGGGGCTTATGTCTGCGCGCCACACAGCCC 957
 Qy 317 LeuThrArgAspPheAlaThrAlaAlaGluLeuArgValLeuAlaAlaProAlaIle 336
 Db 958 CGCACCGCGGACTTCCGCACTGCAGCGCTGAGCTCGGTGTGTGGCGCTCCCGCATC 1017
 Qy 337 SerGlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArg 356
 Db 1018 ACTCAGCGCCCGAGCGCTGTGCGGAGCGGGCGAGCACAGCGGCTGTGCTGTGCCGC 1077
 Qy 357 AlaSerGlyGluProArgProAlaLeuHisTrpLeuHisAspGlyLysLeuProLeuArgPro 376
 Db 1078 GCGTGGGGGAGCGCGGCGGCTGCGCTGCTGCTGCAACAGGGGGCGCTGCGGCC 1137
 Qy 377 AsnGlyArgValLysValGlnGlyGlySerLeuValLysLeuThrGlnLysLeu 396
 Db 1138 AACGGCGCGTCAAGGTCCAGGGCGGCTGGCGAGCTGTGTCTATCACACAGATCGGCTG 1197
 Qy 397 GlnAspAlaGlyTyrTyrGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAla 416
 Db 1198 CAGGACCGCGCTACTACCACTGCTGGCTGAGAACAGCGCGGAAATGGCGTGGCTGCC 1257
 Qy 417 AlaProLeuAlaValValValGluGlyLeuProSerAlaProThrArgValThrAla 436
 Db 1258 GCGTCTGCTGCGCTGTGTGGCGAGGGGCTGCCAGCGCCCCCAGCGGGTCACTGCT 1317
 Qy 437 ThrProLeuSerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGlu 456
 Db 1318 AGCCCACTGAGCAGCTCCGCTGTGTGGTGGCTGGAGCGGCCGAGATGCACAGCGAG 1377
 Qy 457 GlnLysLeuGlyPheSerLeuHisTyrGlnLysAlaArgGlyValAspAsnValGluTyr 476
 Db 1378 CAGATCATCGCTTCTCTCCACTACAGAGGCGACGGGGCAATGGCAATGTGGATAC 1437
 Qy 477 GlnPheAlaValAsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThr 496
 Db 1438 CAGTTTGCAGTGAACACACACACAGAACTACAGGTTCGGGACCTTGGAAACCAACACA 1497
 Qy 497 AspTyrGluPheTyrValValAlaTyrSerGlnLysGlyAlaSerArgThrSerSerPro 516
 Db 1498 GATTATGAGTTCTACGTGTGGCTTACTCCAGCTGGGAGCGACCGCACTCCACCCCA 1557
 Qy 517 AlaLeuValHisThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSer 536
 Db 1558 GCACCTGGTGACACACTGGATGAT-----GGTAGGGCTCTGAACCTCGCAGTGGGCGAGC 1611
 Qy 537 ProAsnProSerAspIleArgValAlaTrpLeuProLeuProSerLeuSerAsnGly 556
 Db 1612 -----TTGGGCTTGGACCAATGGG 1629
 Qy 557 GlnValLeuLysTyrLysLeuGlyTyrGlyLeuGlyLysGluAspGlnValPheSerThr 576
 Db 1630 CAGTGTGAGTACAGATAGATAGTTCGTTGGAAAGGAGATCAGATTTTCTTACT 1689
 Qy 577 GluValProGlyAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyr 596
 Db 1690 GAGTGGAGGAAATGAGACACAGCTTATGCTGAACCTCGCTTCAGCCCAACAGGTGTAT 1749
 Qy 597 ArgValArgLysAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGln 616
 Db 1750 CGAGTAGCGATTTTCGGCTGTACAGCAGCGGCTTCGGGGCCCCCTCCAGTGTGATGCAT 1809
 Qy 617 HisArgThrProGlyValHisAsnGlnSerHisValProPheAlaProAlaGluLeuLys 636
 Db 1810 CACAGACGCCCATGATGCAACACAGAGCATGTCTTTCCTTTTGGCCCTGAGAGTTGAG 1869
 Qy 637 ValArgAlaLysMetGluSerLeuValValSerTrpGlnProProHisProThrGln 656

Db 1870 GTGCGAGCAAGATGAGTCCCTGTCGTGTCATGCGACCCACCCCTCACCCACCCAG 1929
 QY 657 IISerGlyTyrLysLeuTyrTrpGlyGluValGlyThrGluGluAlaAspGlyAsp 676
 Db 1930 ATCTCTGGCTACAACTATTATTTGGCGGAGTGGGGCTGAGAGAGAGCCCATGGCAT 1989
 QY 677 ArgProProGlyGlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLysLys 696
 Db 1990 CGCCTGCCAGGGGGCGGTGAGACACAGGCTTGGGATGGGGCTGTCCGGCTCAAGAG 2049
 QY 697 LysValLysGlnTyrGluLeuThrGlnLeuValProGlyArgProTyrGluValLysLeu 716
 Db 2050 AAAGTGAAGCATATGAGCTGACCCAGCTAGTCCCTGGCCGGCTGTACGAGGTGAAGCTC 2109
 QY 717 ValAlaPheAsnLysHisGluAspGlyTyrAlaAlaValTrpLysGlyLysThrGluLys 736
 Db 2110 GTGGCTTTCAACAACATGCGATGGCTATGCGAGCTGTGGAGGGCGACGAGAG 2169
 QY 737 AlaProThrPro 740
 Db 2170 GCGCCGGCACCA 2181
 RESULT 7
 ABX71435
 ID ABX71435 standard; cDNA; 3453 BP.
 XX
 AC ABX71435;
 XX
 DT 06-MAR-2003 (first entry)
 XX
 DE Human cDNA encoding a novel tyrosine phosphatase, NHP4.
 KW Human; ss; gene; chromosome 15; chromosome 3; tyrosine phosphatase;
 KW Ig super family protein; gene therapy; NHP; novel human protein.
 XX
 OS Homo sapiens.
 PN US6465632-B1.
 XX
 PD 15-OCT-2002.
 XX
 PF 08-JUN-2001; 2001US-0877730.
 XX
 PR 09-JUN-2000; 2000US-210607P.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 XX
 PI Walke DW, Scoville J, Turner CA, Friedrich G, Abuin A;
 PI Zambrowicz B, Sands AT;
 DR WPI: 2003-147071/14.
 DR P-PSDB; ABU54203.
 XX
 PT Novel isolated nucleic acid which encodes a novel human protein that
 PT shares sequence similarity with animal phosphatases, that is useful for
 PT generating antibodies, and as reagents in diagnostic assays
 XX
 PS Claim 1; Column 31-36; 58pp; English.
 CC
 CC The invention relates to an isolated nucleic acid molecule comprising a
 CC nucleotide sequence encoding a novel human protein (NHP) that shares
 CC sequence similarity with animal phosphatases (in particular tyrosine
 CC phosphatases and is a member of the Ig superfamily). The disclosed NHP
 CC polynucleotide sequences (from genes located on either chromosome 15
 CC or chromosome 3) are useful for identifying coding sequences and in the
 CC identification of biologically relevant splice junctions. The NHP
 CC polynucleotide sequences are useful in gene therapy, for detecting
 CC mutant NHPs or inappropriately expressed NHPs, for the diagnosis of
 CC disease, for screening drugs effective in treatment of symptomatic or
 CC phenotypic manifestations of perturbing the normal function of NHP.
 CC Complementary sequences of the NHP polynucleotides used in conjunction
 CC with PCR to screen libraries, isolate clones and prepare cloning and
 CC sequencing templates. Labeled NHP nucleotide probes can be used to screen

CC a human genomic library which is helpful for identifying polymorphisms,
 CC determining the genomic structure of a given locus/allele and designing
 CC diagnostic tests. The probe sequences also have use in defining and
 CC monitoring both drug action and toxicity. Oligonucleotides
 CC complementary to NHPs may encode or act as NHP antisense molecules, or
 CC may be used as part of ribozyme and/or triple helix sequences.
 CC The present sequence encodes an NHP.
 XX
 SQ Sequence 3453 BP; 979 A; 821 C; 769 G; 883 T; 1 other;

Alignment Scores:
 Pred. No.: 5,79e-75 Length: 3453
 Score: 1747.50 Matches: 397
 Percent Similarity: 54.95% Conservative: 169
 Best Local Similarity: 38.54% Mismatches: 381
 Query Match: 26.46% Indels: 83
 DB: 25 Gaps: 18

US-09-754-997A-2 (1-1252) x ABX71435 (1-3453)
 QY 9 GlyLeuLeuValLeuThrPheCysLeuLeuSerAlaArgGlyGluLeuProLeuProGln 28
 Db 43 GGGATGCTGCTCCGGCGCTCTCTGCTCTGCTGCTGCTCAGT-----CCTTGGCCAGGA 96
 QY 29 GluThrThrVal---LysLeuSerCysAspGluGlyProLeuGlnValLysLeuGlyPro 47
 Db 97 GTGTGGTGTCTTAGCGAACTGCTCTTTGTAAAGAACCCACAGGATGTAAGTGTCAAGA 156
 QY 48 GluGlnAlaValValLeuAspCysThrLeuGlyAlaThrAlaAlaGlyProThrArg 67
 Db 157 AAGGACCCAGTCGTTTATAGATTGCCAGCTCCAGCGAGAAAGTT-----CCTATTAAAG 207
 QY 68 ValThrTrpSerLysAspGlyAspThrValLeuGluHisGluAsnLeuHisLeuLeuPro 87
 Db 208 GTCATGTTGTAATAATGGAGCAAAATGCTGTAATAATGAACGATCGAGTCTTTCTT 267
 QY 88 AsnGlySerLeuTrpLeuSerSerProLeuGluGlnGluAspSerAspAspGluGluAla 107
 Db 268 AACGGCTCTTTATACATCAGT-----GAGGTGGAAGCAGCGAGGAGAGCAGTCC 318
 QY 108 LeuArgIleTrpLysValThrGluGlySerTyrSerCysLeuAlaHisSerProLeuGly 127
 Db 319 -----GATGAAGGATTTTATCAGTCTTGGCAATGAACAAATATGGA 360
 QY 128 ValValAlaSerGlnValAlaValValLysLeuAlaThrLeuGluAspPheSerLeuHis 147
 Db 361 GCCATCTTAGTCAAAAGCTCATCTGCTCTTATCACTATTCTGCAATTTGAAGTCAG 420
 QY 148 ProGluSerGlnIleValGluGluAsnGlyThrAlaArgPheGluCysHisThrLysGly 167
 Db 421 CCAATTTCCACTGAGGTCCACGAAGGTGGAGTGTGCTGATTTGCAAGATTTCATCC 480
 QY 168 LeuProAlaProIleIleThrTrpGluLysAspGlnValThrValPro---GluGluPro 185
 Db 481 CACCTCTCTGCGCATATAACATGGAGTTCAATCGGACAACTCTACCTATGACTATGGAC 540
 QY 187 ArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuAspValGlnAspSerAspAla 206
 Db 541 AGGTAACCTGCCCTTACCACAGGAGTATTGCAATCTATGATGTCGCAAGGAGGATTTCT 600
 QY 207 GlySerTyrArgCysValAlaThrAsnSerAlaArgGlnArgPheSerGlnGluAlaSer 226
 Db 601 GGAAATATTCTGTTGTTGCTGCCACTGTAGCCACCGACGACGTAAAGTATGAGGCGCTCG 660
 QY 227 LeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAspValIleVal 246
 Db 661 CTAACGTGTGATT-----CCAGCTAAGGAGTCAAAATCTCTCCACACCAACCAATATA 714
 QY 247 AlaAlaProGluAsnThrThrValValSerGlyGlnAsnValValMetGluCysValAla 266
 Db 715 GCAGGTCCACAGACATTAACCAATCTCTTTCATCAGACTGTAGTATTGGAATGATGCC 774
 QY 267 SerAlaAspProThrProPheValSerTrpValArgGlnAspGlyLysProIleSerThr 286

Db 775 ACAGAAATCCCAACCAATCAATCTTGGAGCGCCTTGATCACAATCCATGATGTC 834
Qy 287 --- AspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSerAlaGlnProArgHis 305
Db 835 TTATAACTCGGGTACTTGGAAATGGTAATCTCATGATACTGATGTCAGGCTACAACAT 894
Qy 306 SerGlyValThrValCysAlaAsnLysProLeuThrArgAspPheAlaThrAlaAla 325
Db 895 GCTGGAGTATATGTTGTCGGGCACCTACCCCTGGCACACGCACTTTACAGTTGCTATG 954
Qy 326 AlaGluLeuArgValLeuAlaProAlaIleSerGlnAlaProGluAlaLeuSerArg 345
Db 955 GCACTTTAACTGATAGTCTCTCTCAATTTGTAATGGCCAGAAATTTAAACAGG 1014
Qy 346 ThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLeu 365
Db 1015 CCTCGAGCTGGCACTGCTCGATTTGTGTGTCAGGCGAAGGAATCCCTCTCCCAAGATG 1074
Qy 366 HisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgValLysValGlnGlyGly 385
Db 1075 TCATGGTTGAAATGGAAGGAAGATACATTCGAATGTTAGAATTTAAATGTATC- 1128
Qy 386 GlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrrGlnCysVal 405
Db 1129 AACAGTAAATGGTAATTAACAGATTAATTCCTGAAGATGATGCTATTTATCAGTGCAT 1188
Qy 406 AlaGluAsnSerAlaGlyThrAlaCysAlaAlaProLeuAlaValValArgGlu 425
Db 1189 GCTGAGNATAGCCAGATCTATTTATCTAGAGCCAGATGACTGTAGTGTGTGTCAGAA 1248
Qy 426 GlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerValLeu 445
Db 1249 GACAGACCCAGTGTCTCTTAATGTATCATGTCTGAAACCATGTCAAGCTCAGCCATCTT 1308
Qy 446 ValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTyrr 465
Db 1309 TTAGCTTGGGAGGCGCCACTTTAATTTACAGAAAGTCAATGCCATTTCTGTACACTAC 1368
Qy 466 GlnLysAlaArgGlyValAspAsnValGluTyrrGlnPheAlaValAsnAspThrThr 485
Db 1369 ATGAAAGCAAGGTTTAATATGAAGATATCAAGTAGTATCGGAAATGACAACT 1428
Qy 486 GluLeuGlnValArgAspLeuGluProAsnThrAspTyrrGluPheTyrrValAlaAlaTyrr 505
Db 1429 CATATATATTATGATGACTTAGAGCTGCCAGCAATATATCTTCTACATGTAGCATAT 1488
Qy 506 SerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThrLeuAspAspVal 525
Db 1489 ATGCCAATGGGAGCCAGCCAGATGTCTGACCATGTGTGACACAGAATATCTTAGAGGATGT 1548
Qy 526 ProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAspIleArgValAla 545
Db 1549 CCCTGAGACCTCTGAAATAGTTTGACAAAGTCGAAGTCCCAAGTCCCACTGATATTCATCTCC 1608
Qy 546 TrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLysTyrrLysIleGluTyrr 565
Db 1609 TGGCTGCCAATCCAGCAATATCGCGGCGGCCAGTGTGTGTGATCGTCTGTCTTTC 1668
Qy 566 GlyLeuGlyLysGluAspGlnValPheSerThrGluValProGlyAsnGluThrGlnLeu 585
Db 1669 CGCTTAAGTACTGAGAAATTCACCAAGTTCTGAGCTCCCGGGGACACGCGATGATC 1728
Qy 586 ThrLeuAsnSerLeuGlnProAsnLysValTyrrArgValArgIleSerAlaGlyThrGly 605
Db 1729 CTTTGGAGGCGCTGAAACCTGACAGTGTCTACTGTGTTCTGGATTAATCTGCTGCCACAGA 1788
Qy 606 AlaGlyTyrrGlyValProSerGlnTrpMetGlnHisArgThrProGlyValHisAsnGln 625
Db 1789 GTGGGCTGGAGAGTATCATGATGACTTCACATAGAGCGCCAAAGCT---ACAAGC 1845
Qy 626 SerHisValProPheAlaProAlaGluLeuLysValArg---AlaLysMetGluSerLeu 644
Db 1845

Db 1846 GTGAAGGCCCTTAAGTCTCCA---GAGTTGCAATTTGGAGCCTCTGAACCTGTACCACCAT 1902
Qy 645 ValValSerTrpGlnProProHisProThr---GlnIleSerGlyTyrrLysLeuTyrr 663
Db 1903 TCTGTGAGGTGGCAGCAAGATGTAGAGGACACAGCTGCTATTTCAGGGCTACAGCTGTAC 1962
Qy 664 TrpGlyGluValGlyThrGluGluAlaAspGlyAspArgProGlyGlyArgGly 683
Db 1963 TACAAGAGAAGAGGCGCAGGAGAAAT--- 1989
Qy 684 AspGlnAlaTrpAspValGlyProValArgLeuLysLysValLysGlnTrpGluLeu 703
Db 1990 ---GGGCCCATTTCTTGATACCAAGACCTACTCTATACTCTC 2031
Qy 704 ThrGlnLeuValProGlyArgProTyrrGluValLysLeuValAlaPheAsnLysHisGlu 723
Db 2032 AGTGGCTTAGACCCCAAGAAATAATCATGTGAGACTCTCTGCTTACAACAACATAGAC 2091
Qy 724 AspGlyTyrrAlaAlaValTrpLysGlyLysThrGluLysAlaProThrProAspLeuPro 743
Db 2092 GATGGCTATCAGGCA---GATCAGACTGTCCAGCACTCCAGGATCGGTG 2136
Qy 744 IleGlnArgGly---ProProLeuProProAlaHisValHisAlaGluSer 759
Db 2137 TCTGTTGATGATGCGCATGTCTCTCCACCAACCACCATCTCTATGGAAGGCT 2196
Qy 760 AsnSerSerThrSerIleTrpLeuArgTrpLysLysProAspPheThrThrValLysIle 779
Db 2197 AACACCTCATCTTCCATCTTCTGCACTGAGGAGGCTGCTCATTCACCGCTGCACAAATC 2256
Qy 780 ValAsnTyrrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTyrr 799
Db 2257 ATTAACATACACCATCGCTGTAACTCTGTGGCTGCAGAACTCTTCTTGGTCTGTATC 2316
Qy 800 TyrrThrSerSerGlyGluAspIleLeuIleGlyLysLeuLysProPheThrThrTyrrGlu 819
Db 2317 CTTCAACATCAGAAACTCAGATGTGTGTTCAAGTCTAGAACTAGAACCAACCAATACGAA 2376
Qy 820 PheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArg 839
Db 2377 TTGCGCTTCGATTACATGTGGATCAGCTTTCAGTCTTCCAGCTCTGGAGCCTGTAGTCTACC 2436
Qy 840 SerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThr 859
Db 2437 TCTACTCTTCCAGAGCAGCAGCAGCCACCACTGTTGGAGTAAAGTGAATTAATAGAG 2496
Qy 860 ProSerThrValArgLeuHisTrpCysProThrGluProAsnGlyGluIleValGlu 879
Db 2497 GATGACACTGCCCTGCTGTTCTTGGAAACCCCTGATGGCCCAAGACAGTTGTGACCCGC 2556
Qy 880 TyrrLeuIleLeuTyrrSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuThr 899
Db 2557 TATACTATCTTATATGATCATCTAGGAAGGCTGGATTTCAGGAGAGTGGCAGGCTTACAC 2616
Qy 900 ThrGluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrrPhe 919
Db 2617 CGTGAAGGGCAATAACCATGGCTTGTGTAGAAAACCTTGTAGCAGGAAATGTGTACATT 2676
Qy 920 PheLysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspVal 939
Db 2677 GTCAGATATCTGCTCCCAATGAGGTGGGAGGAGGACCCCTTTTCAATTTCTGTGGAGCTG 2736
Qy 940 IleThrLeuGlnGluThrPheSerAspSer--- 949
Db 2737 GCAGTACTTCCAAAGGAAACCTCTGAATCAATCAGAGGCCCAAGCGTTTAGATTCTGCT 2796
Qy 950 ---LeuAspValHisAlaValThrGlyIleIle 959
Db 2797 GATGCCAAAGTTTATTTCAGGATATTACATCTGCACCAAAAATCAATGATGCTGCATTGCT 2856
Qy 960 ValGlyValCysLeuGlyLeuLeuCysLeuAlaCysMetCysAlaGlyLeuArgGln 979
Db 2857 GTAGGTGTTGGCATAGCTTGCATCTCATCTGTGTTCTCTCTCTGATATACCGA 2916

Db 835 TTTAACTACTCGGGTACTTGGAAATGGTAATCTCATGATATCTGATGTGCGGCTACCAACAT 894
Qy 306 SerGlyValTyrValCysArgAlaAsnLysProLeuThrArgAspPheAlaThrAlaAla 325
Db 895 GCTGGAGTATATGTTGTGCGGCCACTACCCCTGGCACACCACTTTACAGTTGCTATG 954
Qy 326 AlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerArg 345
Db 955 GCACCTTTAACTGTATTAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1014
Qy 346 ThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLeu 365
Db 1015 CCTCGAGCTGGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1074
Qy 366 HisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgValLysValGlnGlyGly 385
Db 1075 TCATGGTTGAAATGGAAGGAGATACATTCGATGGTAGAATTAATAATGTAC----- 1128
Qy 386 GlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrTyrGlnCysVal 405
Db 1129 AACAGTAAATTTGGTAATTAACAGATTTATTCCTCAAGATGATGCTATTTATCAGTGCATG 1188
Qy 406 AlaGluAsnSerAlaGlyThrAlaCysAlaAlaAlaProLeuAlaValValAlaArgGlu 425
Db 1189 GCTGAGAATAGCCAGGATCTATTTTATCTAGAGCCAGACTGACTGTAGTGTGTCAGAA 1248
Qy 426 GlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerValLeu 445
Db 1249 GACAGACCCAGTCTCCTATATATGATATGATCTGAAACCATGTCAGCTGAGCTGCTTCT 1308
Qy 446 ValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTyr 465
Db 1309 TTAGCCTGGGAGAGGCCACTTTATATTACAGAAAGTCAATGCTTCTGCTATCTGTACACTAC 1368
Qy 466 GlnLysAlaArgGlyValAspAsnValGluTyrGlnPheAlaValAlaAsnAsnAspThrThr 485
Db 1369 ATGAAAGCAGAGGTTTAAATATGAAAGATATCAAGTAGTCAATCGAAATGACACAACT 1428
Qy 486 GluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPheTyrValValAlaTyr 505
Db 1429 CATTATATTATGATGACTTAGAGCCTGCGCAGCAATTTATCTTTACATTTAGCATAT 1488
Qy 506 SerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThrLeuAspAspVal 525
Db 1489 ATGCCAATGGGAGCCAGCCAGATGCTCTGACCATGTGCACAGAATACTCTAGAGGATGTT 1548
Qy 526 ProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAspIleArgValAla 545
Db 1549 CCCCTGAGACCTCTCTGAAATTTAGTTTGACAAGTCGAAGTCCCACTGATATTTCTCATCTCC 1608
Qy 546 TrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLysTyrLysIleGluTyr 565
Db 1609 TGGCTGCCAATCCAGCCAAATATCGCGGGGCCAAGTGTGCTGTATCGCTGTCTTCTTC 1668
Qy 566 GlyLeuGlyLysGluAspGlnValPheSerThrGluValProGlyAsnGluThrGlnLeu 585
Db 1669 CGCCTAAGTACTGAGAAATCAATCAAGTCTCGAGCTCCCGGGGCCACCGCATGAGTAC 1728
Qy 586 ThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgIleSerAlaGlyThrGly 605
Db 1729 CTTTGGAGGCCCTGAAACCTGACAGTGTCTACTGTTCCGATTTACTGCTGCCACAGA 1788
Qy 606 AlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrProGlyValHisAsnGln 625
Db 1789 GTGGGGCTGGGAGAGTCAATCAGTATGGACTTTCATAGGAGCCGCCAAAGCT---ACAAGC 1845
Qy 626 SerHisValProPheAlaProAlaGluLeuLysValArg---AlaLysMetGluSerLeu 644
Db 1846 GTGAAGCCCTTAAGTCTCA---GAGTTCATTTGGAGGCTCTGAACTGTACACCAT 1902
Qy 645 ValValSerTrpGlnProProHisProThr---GlnIleSerGlyTyrLysLeuTyr 663

Db 1903 TCTGTGAGTGGGAGCAAGATGTAGAGGACACAGCTGCTATTTCAGGGCTACAGCTGTAC 1962
Qy 664 TrpGlyGluValGlyThrGluGluAlaAspGlyAspArgProGlyGlyArgGly 683
Db 1963 TACAAGGAAGAGGGCAGCAGAGAGAT----- 1989
Qy 684 AspGlnAlaTrpAspValGlyProValArgLeuLysLysValLysGlnTyrGluLeu 703
Db 1990 -----GGGCCCATTTTCTTGGATACCAAGGACCTACTCTATATCTCTC 2031
Qy 704 ThrGlnLeuValProGlyArgProTyrGluValLysLeuAlaPheAsnLysHisGlu 723
Db 2032 AGTGGCTTAGACCCCGAAGAAATATCATGTGAGACTCTGTGCTTACCAACATAGAC 2091
Qy 724 AspGlyTyrAlaAlaValTrpLysGlyLysThrGluLysAlaProThrProAspLeuPro 743
Db 2092 GATGGCTATCAGGCA-----GATCAGACTGTTCAGCACTCCAGGATCGGTG 2136
Qy 744 IleGlnArgGly-----ProProLeuProProAlaHisValHisAlaGluSer 759
Db 2137 TCTGTTTCGTGATGCGCAGTCCCTCCTCCACCAACCCACCATCTCTATGCGAAGCT 2196
Qy 760 AsnSerSerThrSerIleTrpLeuArgTrpLysLysProAspPheThrThrValLysIle 779
Db 2197 AACCTCTCATCTTCCATCTCTCTGCACTGGAGAGGCCCTGCAATTCACCGCTGCACAAATC 2256
Qy 780 ValAsnTyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTyr 799
Db 2257 ATTAACATACACCATCGCTGTAATCTCTGTGGCTGCAGATGCTTCTTGGTCTGTAC 2316
Qy 800 TyrThrSerSerGlyGluAspIleLeuIleGlyLeuLysProPheThrLysTyrGlu 819
Db 2317 CTTCAACATCATGAAACTCATGTTGTTCAAGGCTTAGAACCAACCAACCAATACGAA 2376
Qy 820 PheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArg 839
Db 2377 TTTGCGGTCGATTAATCATGATGATCAGCTTCCAGTCTTCCAGGCCCTGTAGTCTACCAT 2436
Qy 840 SerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThr 859
Db 2437 TCTACTCTTCCAGAACACCCAGCCGCCACCCAGTTGGAGTAAAGTGCATTAATAGAG 2496
Qy 860 ProSerThrValArgLeuHisTrpCysProProThrGluProAsnGlyGluIleValGlu 879
Db 2497 GATGACACTGCCCTGCTTTCTGAAACCCCTGATGGGCCAGAAACAGTTGTGACCCGC 2556
Qy 880 TyrLeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThr 899
Db 2557 TATACTATCTTATGTCATCTAGGAAGGCTGGATTCGAGGAGAGTGGCAGGCTTTCACAC 2616
Qy 900 ThrGluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhe 919
Db 2617 CGTGAAGGGGCAATACCATGCTTGTCTAGAAAACCTTGTAGCAGGAATGTGTACATT 2676
Qy 920 PheLysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspVal 939
Db 2677 GTCAAGATATCTGATCAATCAGGTGGGAGGAGGCCCTTTTCAAATTTCTGTGGAGCTG 2736
Qy 940 IleThrLeuGlnGluThrPheSerAspSer----- 949
Db 2737 GCAGTACTTCCAAAGGAACCACTCTGAATCAATTCAGAGGCCCAAGCGTTTAGATTCTGCT 2796
Qy 950 -----LeuAspValHisAlaValThrGlyIleIle 959
Db 2797 GATGCCAAAGTTTATTCAGATATTTACCATCTGGACCAAAATCAATGATGGCATGCT 2856
Qy 960 ValGlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGln 979
Db 2857 GTAGGTGTTGGCATGACCTTGACCTGATCTCTCTGTTCTCTCATCTTGTATATACCGA 2916
Qy 980 SerSerHisArgGluAla-----LeuPro 987
Db 2917 AGTAAAGCCAGGAATCATCTGCTCTCCAAAGCCGACAGAACTGGAATCAACAGTTTACCT 2976

Qy 988 GlyLeuSerSerSerGlyThrProGlyValAsn 997
 Db 2977 CGTACCAGTCCCTCTAGTAGTGGAAAT 3006
 RESULT 9
 ID ABX71447
 AC ABX71447 standard; cDNA; 3874 BP.
 XX ABX71447;
 DT 06-MAR-2003 (first entry)
 XX Human cDNA encoding a novel tyrosine phosphatase.
 XX Human; ss; gene; chromosome 15; chromosome 3; tyrosine phosphatase;
 KW ig super family protein; gene therapy; NHP; novel human protein.
 XX Homo sapiens.
 OS US6465632-B1.
 PN 15-OCT-2002.
 XX 08-JUN-2001; 2001US-0877730.
 XX 09-JUN-2000; 2000US-210607P.
 XX (LEXI-) LEXICON GENETICS INC.
 XX Walke DW, Scoville J, Turner CA, Friedrich G, Abuin A;
 PI Zambrowicz B, Sands AT;
 XX WPI; 2003-147071/14.
 DR Novel isolated nucleic acid which encodes a novel human protein that
 PT shares sequence similarity with animal phosphatases, that is useful for
 PT generating antibodies, and as reagents in diagnostic assays -
 XX Disclosure; Column 107-112; 58pp; English.
 CC The invention relates to an isolated nucleic acid molecule comprising a
 CC nucleotide sequence encoding a novel human protein (NHP) that shares
 CC sequence similarity with animal phosphatases (in particular tyrosine
 CC phosphatases and is a member of the Ig superfamily). The disclosed NHP
 CC polynucleotide sequences (from genes located on either chromosome 15
 CC or chromosome 3) are useful for identifying coding sequences and in the
 CC identification of biologically relevant splice junctions. The NHP
 CC polynucleotide sequences are useful in gene therapy, for detecting
 CC mutant NHPs or inappropriately expressed NHPs, for the diagnosis of
 CC disease, for screening drugs effective in treatment of symptomatic or
 CC phenotypic manifestations of perturbing the normal function of NH.
 CC Complementary sequences of the NHP polynucleotides used in conjunction
 CC with PCR to screen libraries, isolate clones and prepare cloning and
 CC sequencing templates. Labeled NHP nucleotide probes can be used to screen
 CC a human genomic library which is helpful for identifying polymorphisms,
 CC determining the genomic structure of a given locus/allele and designing
 CC diagnostic tests. The probe sequences also have use in defining and
 CC monitoring both drug action and toxicity. Oligonucleotides
 CC complementary to NHPs may encode or act as NHP antisense molecules, or
 CC may be used as part of ribozyme and/or triple helix sequences.
 CC The present sequence encodes an NHP.
 XX
 SQ Sequence 3874 BP; 1077 A; 938 C; 897 G; 960 T; 2 other;

Alignment Scores:

Pred. No.:	7,318-75	Length:	3874
Score:	1746.50	Matches:	397
Percent Similarity:	54.95%	Conservative:	169
Best Local Similarity:	38.54%	Mismatches:	381
Query Match:	26.45%	Indels:	83
DB:	25	Gaps:	18

US-09-754-997A-2 (1-1252) x ABX71447 (1-3874)
 Qy 9 GlyLeuValLeuThrPheCysLeuSerAlaArgGlyGluLeuProLeuProGln 28
 Db 233 GGGAGTGTCTCCGGCGCTCTGCTCTGCTGCTGCTCAGT-----CCTTGGCAGA 286
 Qy 29 GluThrThrVal---LysLeuSerCysAspGluGlyProLeuGlnValLeuLeuGlyPro 47
 Db 287 GTGTGTGTCTTTAGCGAAGCTGCTTTTGTAAAGAACACACAGATGTAACTGTCAACA 346
 Qy 48 GluGlnAlaValValLeuAspCysThrLeuGlyAlaThrAlaAlaGlyProThrArg 67
 Db 347 AAGGACCCAGTGTGTTTATAGTTGCCAGGCTCAGGAGAAGTT-----CCTATTAG 397
 Qy 68 ValThrTrpSerLysAspGlyAspThrValLeuGluHisGluAsnLeuHisLeuLeuPro 87
 Db 398 GTCACATGTTGAAAATGAGCAAAAATGCTGTGAAAATAAACAGATCAGGTCTTTCT 457
 Qy 88 AsnGlySerLeuTrpLeuSerProLeuGluGlnGluAspSerAspAspGluAla 107
 Db 458 AACGGCTCTTTATACATCAGT-----GAGGTGAAGGCGAGGCGAGAGCAGTC 508
 Qy 108 LeuArgIleTrpLysValThrGluGlySerTyrSerCysLeuAlaHisSerProLeuGly 127
 Db 509 -----GATGAAGGATTTTATCAGTGTCTGCAATGAAACAATATGA 550
 Qy 128 ValValAlaSerGlnValAlaValLysLeuAlaThrLeuGluAspPheSerLeuHis 147
 Db 551 GCCATTCTTAGTCAAAAGCTCATCTTGCTTATCAACTATTTCTGCATTTGAGTCCAG 610
 Qy 148 ProGluSerGlnIleValGluGluAsnGlyThrAlaArgPheGluCysHisThrLysGly 167
 Db 611 CCAATTTCCACTGAGGTCCACGAAGGTGGAGTTGCTCGATTTGCATGCAAGATTTTCATCC 670
 Qy 168 LeuProAlaProIleIleThrTrpGluLysAspGlnValThrValPro---GluGluPro 186
 Db 671 CACCTCTCTGCACTAATACATGGAGTTCAATCGGCAACTCTACCTATGATATGCGAC 730
 Qy 187 ArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuAspValGlnAspSerAspAla 206
 Db 731 AGATACTGCCCTACCAACAGGATTTGCAATCTATGATGTGAGCCAAAGGATCT 790
 Qy 207 GlySerTyrArgCysValAlaThrAsnSerAlaArgGlnArgPheSerGlnGluAlaSer 226
 Db 791 GGAATATGTTGTTGTTGCTGCCACTGTAGCCACCGACGTAATAAGTATGGAGGCTCG 850
 Qy 227 LeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAspValIleVal 246
 Db 851 CTAACGTGTGATT-----CCAGCTAAGAGTCAAAATCTTCCACACACCAACCAATATA 904
 Qy 247 AlaAlaProGluAsnThrThrValValSerGlyGlnAsnValValMetGluCysValAla 266
 Db 905 GCAGGTCCACAGAACATAAACACATCTCTTCATCAGACTGTAGTTTGGATGATGATGCC 964
 Qy 267 SerAlaAspProThrProPheValSerTrpValArgGlnAspGlyLysProIleSerThr 286
 Db 965 ACAGGAATCCCAACCAATCAITTTCTTGAGCGCCTTGCATCACAATCCATTGATGTC 1024
 Qy 287 ---AspValIleValLeuGlyArgThrAsnLeuLeuAlaSerAlaGlnProArgHis 305
 Db 1025 TTTAATACTCGGTACTTGGAAATGGTAATCTCATGATATCTGATGTCAGCGTACAACAT 1084
 Qy 306 SerGlyValTrpValCysArgAlaAsnLysProLeuThrArgAspPheAlaThrAlaAla 325
 Db 1085 GCTGGAGTATATGTTGTGCGGCACCTACCCCTGGCACAGCAACTTTACAGTTGCTATG 1144
 Qy 326 AlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerArg 345
 Db 1145 GCAACTTTAAGTATGATGCTCTCTCTTCTTGTGATGGCCAGAAAGTTTAAACAGG 1204
 Qy 346 ThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLeu 365
 Db 1205 CCTCGAGCTGGCACTGCTCGATTGTTGTGTGTCAGGACGAAGAAATCCCTCTCTCCCAAGATG 1264

QY 366 HisTrpLeuHisaspGlyIleProLeuArgProAsnGlyArgValIysValGlnGlyGly 385
DB 1265 TCATGGTTGAAATGGAAGCAATACATTCGAATGCTAGAAATTAATAATGTAC----- 1318
QY 386 GlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrGlnCysVal 405
DB 1319 AACAGTAAATGGTAATTAACAGATTAATTCCTGAAGATGATCTATTATCATAGTGCATG 1378
QY 406 AlaGlnAsnSerAlaGlyThrAlaCysAlaAlaProLeuAlaValValValArgGlu 425
DB 1379 GCTGAGAATAGCAAGGATCTATTATCTAGAGCCAGACTGACTGATGATGCAGAA 1438
QY 426 GlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerValLeu 445
DB 1439 GACAGACCCAGTGTCTCCCTAATATGATCATGCTGAAACCAATGTCAAGCTCAGCCATCTT 1498
QY 446 ValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTyr 465
DB 1499 TTAGCCTGGGAGAGGCCACTTTATATTTCAGACAAGTCAATGCCATTCTGTACACTAC 1558
QY 466 GlnIysAlaArgGlyValAspAsnValGluTyrGlnPheAlaValAsnAspThrThr 485
DB 1559 ATGAAGCAGAGGTTTAATATGAAGATATCAAGTAGTCAATCGAATGACACACT 1618
QY 486 GluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPheTyrValValAlaTyr 505
DB 1619 CATATATTATTGATGACTTAGAGCCTGCCAGCAATATATATCTTCTACATGTAGCATAT 1678
QY 506 SerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThrLeuAspAspVal 525
DB 1679 ATGCCAATGGAGCCAGCCAGATGTCTGACCATGTGTGACACAGAAATCTCTAGAGGATGT 1738
QY 526 ProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAspIleArgValAla 545
DB 1739 CCCTGAGACCTCTGAAATAGTTTGACAGTCAAGTCCAGTCCACTGATATCTCATCTCC 1798
QY 546 TrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuIysTyrIleGluTyr 565
DB 1799 TGGCTGCCAATCCAGCCAAATATCGCGGGGCCAAGTGTGTGTATCGTCTGTCTTCTTC 1858
QY 566 GlyLeuGlyIysGluAspGlnValPheSerThrGluValProGlyAsnGluThrGlnLeu 585
DB 1859 CGCTAAGTACTGAGAAATCAATCCAAAGTTCTGAGCTCCCGGGACCACGATGATGATC 1918
QY 586 ThrLeuAsnSerLeuGlnProAsnValValTyrArgValAlaGlySerAlaGlyThrGly 605
DB 1919 CTTTGGAGCCCTTAAGTCTCA--GAGTTGCAATTTGGAGCCTCTGAACTGTACACCANT 1978
QY 606 AlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrProGlyValHisAsnGln 625
DB 1979 GTGGGCTGGAGATCATCAGTATGCACTTCACATAGGAGCCCAAGCT--ACAAGC 2035
QY 626 SerHisValProPheAlaProAlaGluLeuIysValArg--AlaIysMetGluSerLeu 644
DB 2036 GTGAAGCCCTTAAGTCTCA--GAGTTGCAATTTGGAGCCTCTGAACTGTACACCANT 2092
QY 645 ValValSerTrpGlnProProHisProThr---GlnIleSerGlyTyrIysLeuTyr 663
DB 2093 TCTGTAGGTGGCAGCAAGATGTAGAGGACACAGCTCTATTCAGGCTTCAAGCTGTATC 2152
QY 664 TrpGlyGluValGlyThrGluGluGluAlaAspGlyAspArgProGlyGlyArgGly 683
DB 2153 TACAAGGAGAGGGCAGGAGAAAT----- 2179
QY 684 AspGlnAlaTrpAspValGlyProValArgLeuIysLysValIysValIysGlnIleGlu 703
DB 2180 -----GGGCCCATTTCTTGGATACCAAGGACCTACTCTATACTACTCTC 2221
QY 704 ThrGlnLeuValProGlyArgProTyrGluValIysLeuValAlaPheAsnLysHisGlu 723
DB 2222 AGTGGCTTAGACCCCAAGAAATAATCATGTGAGACTCTCGGCTTACAAACAACATAGAC 2281

QY 724 AspGlyTyrAlaAlaValTrpIysGlyIysThrGluIysAlaProThrProAspLeuPro 743
DB 2282 GATGGCTATCAGGCA-----GATCAGACTGTCCAGCACTCCAGGATCGGTG 2326
QY 744 IleGlnArgGly-----ProLeuProProAlaHisValHisAlaGluSer 759
DB 2327 TCTGTTGCTGATCGCATGCTCTCTCCACACACCCACCATCTCTATGCGAAGCT 2386
QY 760 AsnSerSerThrSerIleTrpLeuArgTrpIysLysProAspPheThrThrValIysIle 779
DB 2387 AACACCTCATCTCCATCTTCTGCACTGGAGAGGCTGCATTACCCGCTGCACAAATC 2446
QY 780 ValAsnTyrThrValArgPheGlyProTrpGlyIysArgAsnAlaSerLeuValThrTyr 799
DB 2447 ATTAACATACCATCCGCTGTAATCTGTGGCTGCAGATGCTTCTTGGTCTGTATC 2506
QY 800 TyrThrSerSerGlyGluAspIleLeuIleGlyIysLeuIysProPheThrIysTyrGlu 819
DB 2507 CTTCAACATACAGAACTCAGATGTGGTTCAGGTCTAGAACCAACCAACCAATACGAA 2566
QY 820 PheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArg 839
DB 2567 TTGGCGTTCCGATTACATGTCATCAGCTTTCAGTCTTGGAGCCCTGTAGTCTACCAT 2626
QY 840 SerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThr 859
DB 2627 TCTACTCTTCCAGAACACACAGCCACCATGTTGGAGTAAAGTGCATTAATAGAG 2686
QY 860 ProSerThrValArgLeuHisTrpCysProThrGluProAsnGlyGluIleValGlu 879
DB 2687 GATGACACTGCCCTGTGTTCTGGAAACCCCTGTATGGCCCAAGAACAGTTGACCCGC 2746
QY 880 TyrLeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThr 899
DB 2747 TATACTATCTATATGCACTAGGAAGGCTGATTCAGGAGAGTGGCAGGCTTTACAC 2806
QY 900 ThrGluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhe 919
DB 2807 CGTGAAGGGGCAATACCATGCTTGGCTAGAAAACCTTGTGTAGCAGGAAATGTGTACAT 2866
QY 920 PheIysMetGlyValArgThrGluValGlyProGlyProPheSerArgLeuGlnAspVal 939
DB 2867 GTCAAGATATCTGCATCCCAATGAGGTGGGAGGAGGACCTTTTCAAATTTCTGTGAGCTG 2926
QY 940 IleThrLeuGlnGluThrPheSerAspSer----- 949
DB 2927 GCAGTACTTCCAAAGGAACCTCTGAATCAAATCAGAGCCCAAGCGTTTAGATTCTGCT 2986
QY 950 -----LeuAspValHisAlaValThrGlyIleIle 959
DB 2987 GATGCCAAAGTTTATTTCAGATATTACCATCTGCACCAAAAATCAATGACTGCATTTGCT 3046
QY 960 ValGlyValCysLeuLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGln 979
DB 3047 GTAGGTGTTGGCATACGCTTGACCTGCATCTCTCATCTGTCTCTCATCTTGTATATACCGA 3106
QY 980 SerSerHisArgGluAla-----LeuPro 987
DB 3107 AGTAAGCCAGGAATCATCTGCTTCCAAGCGGACAGAAATGGAATCAACAGTTTACCT 3166
QY 988 GlyLeuSerSerSerGlyThrProGlyAsn 997
DB 3167 CGTACCAGTGCCTCTCTAGCTAGTGAAT 3196

RESULT 10

ABX71437

ID ABX71437 standard; cDNA; 2976 BP.

XX AC ABX71437;

XX DT 06-MAR-2003 (first entry)

XX DE Human cDNA encoding a novel tyrosine phosphatase, NHP6.

CC nucleotide sequence encoding a novel human protein (NHP) that shares
 CC sequence similarity with animal phosphatases (in particular tyrosine
 CC phosphatases and is a member of the Ig superfamily). The disclosed NHP
 CC polynucleotide sequences (from genes located on either chromosome 15
 CC or chromosome 3) are useful for identifying coding sequences and in the
 CC identification of biologically relevant splice junctions. The NHP
 CC polynucleotide sequences are useful in gene therapy, for detecting
 CC mutant NHPs or inappropriately expressed NHPs, for the diagnosis of
 CC disease, for screening drugs effective in treatment of symptomatic or
 CC phenotypic manifestations of perturbing the normal function of NHP.
 CC Complementary sequences of the NHP polynucleotides used in conjunction
 CC with PCR to screen libraries, isolate clones and prepare cloning and
 CC sequencing templates. Labeled NHP nucleotide probes can be used to screen
 CC a human genomic library which is helpful for identifying polymorphisms,
 CC determining the genomic structure of a given locus/allele and designing
 CC diagnostic tests. The probe sequences also have use in defining and
 CC monitoring both drug action and toxicity. Oligonucleotides
 CC complementary to NHPs may encode or act as NHP antisense molecules, or
 CC may be used as part of ribozyme and/or triple helix sequences.
 CC The present sequence encodes an NHP.

SQ Sequence 3219 BP; 927 A; 757 C; 706 G; 828 T; 1 other;

Alignment Scores:

Pred. No.:	2,61e-72	Length:	3219
Score:	1691.50	Matches:	375
Percent Similarity:	55.79%	Conservative:	160
Best Local Similarity:	29.10%	Mismatches:	347
Query Match:	25.61%	Indels:	77
DB:	25	Gaps:	15

US-09-754-997A-2 (1-1252) x ABX71440 (1-3219)

QY	79	GLUHSGLUASenLeuHisSerLeuProAnglySerLeuTrpLeuSerSerProLeuGlu	98
DB	7	GAATAATNACGATCGAGGTTCTTTCTAAACGGCTCTTTATACATCAGT-----GAG	57
QY	99	GLNGluAspSerAspGluAlaLeuArglleTrpLysValThrGluGlySerTyr	118
DB	58	GTGGAGGCGCGGAGGAGGAGTCC-----GATGAGGATTTTAT	99
QY	119	SerCysLeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValLysLeu	138
DB	100	CAGTCTTGCGCAATGAACAATATGAGCCATTCTTAGTCAAAAGCTCATCTTCCTTA	159
QY	139	AlaThrLeuGluAspPheSerLeuHisProGluSerGlnlleValGluGluAsnGlyThr	158
DB	160	TCAACTATTCTGCAATTTGAAGTCAGGCCAATTTCCACTGAGGTCACGAAAGGTGAGTT	219
QY	159	AlaArgPheGluCysHisThrLysGlyLeuProAlaProIlelleThrTrpGluLysAsp	178
DB	220	GCTCGATTTGCGATGCAAGATTTCATCCACCCTCTCGAGTCATTAACATGGAGTTCAAT	279
QY	179	GlnValThrValPro---GluGluProArgLeuIleThrLeuProLysTrpLeuGln	197
DB	280	CGGACAACCTCTACCTATGACTATGACAGGATAACTGCCCTACCAACAGGAGTATTGAG	339
QY	198	IleLeuAspValGlnAspSerAspAlaGlySerTyrArgCysValAlaThrAsnSerAla	217
DB	340	ATCTATGATGTGAGCCAAAGGAGTTCTGGAAATTTATCGTTGTTGCTGCCACTGTAGCC	399
QY	218	ArgGlnArgPheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAla	237
DB	400	CACCGACGTPAAAGATGAGGCGCTCGCTAACTGTGATT-----CCAGCTAAGAGTGCA	453
QY	238	ThrArgGlyGlnAspValIleValAlaAlaProGluAsnThrThrValValSerGly	257
DB	454	AAATCTCTCCACACACCAACCAATATGAGGTCAGGTCACAGAACATAACACATCTCTTCAT	513
QY	258	GlnAsnValValMetGluCysValAlaSerAlaAspProThrProPheValSerTrpVal	277
DB	514	CAGACTGTAGTTTGGATGTCATGGCCACAGGAAATCCCAACCAATCATTTCTTTGGAGC	573

QY	278	ArgGlnAspGlyLysProLysSerThr---AspValIleValLeuGlyArgThrAsnLeu	296
DB	574	CGCCTTGATCACAATCCATTCATGCTTTTAATATCTCGGCTACTTGGAAATGTAATCTC	633
QY	297	LeuIleAlaSerAlaGlnProArgHisSerGlyValTyrValCysArgAlaAsnLysPro	316
DB	634	ATGATATCTCATGTCAGGCTACCAACATGCTGGAGTATATGTTGTGGGCGCACTACCCCT	693
QY	317	LeuThrArgAspPheAlaThrAlaAlaLeuArgValLeuAlaAlaProAlaIle	336
DB	694	GGCACACGCAACTTACAGTTGCTATGGCACTTTAACTGCTATTAGCTCCTCTTCATTT	753
QY	337	SerGlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArg	356
DB	754	GTTGAATGGCCAGAAAGTTTAAACAGGCTCGAGCTGGCACCTGCTCGATTTGTGTGTCAG	813
QY	357	AlaSerGlyCluProArgProAlaLeuHisTrpLeuHisAspGlylleProLeuArgPro	376
DB	814	GCAAGGAATCCCTCTCCCAAGATGTCATGTTGAAATATGGAAGGAAGATCATTCG	873
QY	377	AsnGlyArgValLysValGlnGlyGlyGlySerLeuValIleThrGlnlleGlyLeu	396
DB	874	ANTGTTAGATTAATAATGTAC-----AACAGTAAATTTGTAATTAACACAGATTATTCCT	927
QY	397	GlnAspAlaGlyTyrTyrGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAla	416
DB	928	GAAGATGATGCTATTTATCAGTGCATGCTGAGATAGCAAGGATCTATTATTATCTAGA	987
QY	417	AlaProLeuAlaValValArgGluGlyLeuProSerAlaProThrArgValThrAla	436
DB	988	GCCAGACTGCTGTAGTGTGTGAGAGACAGACCCAGTCTCCCTATAATATGAAGATAT	1047
QY	437	ThrProLeuSerSerSerValLeuAlaTrpGluArgProGluLeuHisSerGlu	456
DB	1048	GAACACATGTCAGCTCAGCCATCTTTTAGCTGGAGAGGCCACTTTTATAATTCAGAC	1107
QY	457	GlnlleGlyPheSerLeuHisTyrGlnLysAlaArgGlyValAspAsnValGlyTyr	476
DB	1108	AAAGTCATTGCTATTCTGTACACTACATGAAAGCAGAGGTTTAAATAATGAAGAGTAT	1167
QY	477	GlnPheAlaValAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThr	496
DB	1168	CAAGTAGTCATCGAAATGACACAACTCATTTATTTATGATGACTTAGAGCTGCCAGC	1227
QY	497	AspTyrGluPheTyrValValAlaTyrSerGlnLeuGlyAlaSerArgThrSerSerPro	516
DB	1228	AATTATCTTTTACATTTGATGATATATGCAATGGAGGCCAGCCAGATGCTCTGACCAT	1287
QY	517	AlaLeuValHisThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSer	536
DB	1288	GTGACACAGAATACTCTAGAGGATGTTCCCTCGAGACTCTCTGAAATTTAGTTGACAAGT	1347
QY	537	ProAsnProSerAspIleArgValAlaTrpLeuProLeuProLeuProSerSerLeuSerAsnGly	556
DB	1348	CGAAGTCCCACTGATATTTCTCATCTCTGCTGCTGCAATCCAGCCAAATATCGCGGGGCG	1407
QY	557	GlnValLeuLysTyrLysIleGluTyrGlyLysGluAspGlnValPheSerThr	576
DB	1408	CAAGTGTGCTGATTCGCTGCTCTTCGCTTAAGTACTGAGAAATTCATCAAGTTCTG	1467
QY	577	GluValProGlyAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyr	596
DB	1468	GAGCTCCCGGGGACCCAGCATGAGTACCTTTTGGAGGCCCTGAAACCTGACAGTGTCTAC	1527
QY	597	ArgValArgIleSerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGln	616
DB	1528	CTGTTTCGATTTACTGCTGCCACCAAGTGGGCTGGGAGAGTCACTAGTAGGACTTCA	1587
QY	617	HisArgThrProGlyValHisAsnGlnSerHisValProPheAlaProAlaGluLeuLys	636
DB	1588	CATAGGACGCCCAAGCT---ACAAGCGTGAAGGCCCTAAGTCTCCA---GAGTTGAT	1641
QY	637	ValArg---AlaLysMetGluSerLeuValSerTrpGlnProProProHisProThr	655

Db	1642	TTGAGGCTCTGAAC	TGAC	CAACATTTCTGTGAGG	GGCAGC	GAAGATGTAGAGG	ACACA	1701
Qy	656	--Gln	IleSerGly	Tyr	Leu	Tyr	Leu	674
Db	1702	GCTGCTATT	CAGGCTAC	AGCTGTACTAC	AGGAAGG	CAGCAGGAG	AAT	1755
Qy	675	GlyAsp	ArgPro	ProGly	Gly	ArgGly	AspGln	694
Db	1756	-----	-----	-----	-----	-----	-----	1770
Qy	695	Lys	Lys	Lys	Val	Lys	Gln	714
Db	1771	GATACCA	AGGACCTAC	TCTATCT	CAGTGGCTT	GAGCCCGA	AGAAATAT	1830
Qy	715	Lys	Leu	Val	Ala	Phe	Asn	734
Db	1831	AGACTCT	GGCTTAC	ACAACAT	AGATGGCTAT	CAGCA	-----	1875
Qy	735	Glu	Lys	Ala	Pro	Thr	Pro	750
Db	1876	CAGACTG	TCAGCACT	CCAGGAT	GGTGTCT	GTGATCG	CATGCTCC	1935
Qy	751	Pro	Pro	Ala	His	Val	His	770
Db	1936	CCACCC	CACCATTCT	TATG	CGAAGCTT	AAACACCT	CATCTTCC	1995
Qy	771	Lys	Pro	Asp	Phe	Thr	Val	790
Db	1996	AGGCTGC	ATTCAC	CGCTGC	ACAAATCAT	TAACTAC	ACCATCCG	2055
Qy	791	Leu	Arg	Asn	Ala	Ser	Leu	810
Db	2056	CTGCAGA	ATGCTTCT	TTGGTTCT	GTACCTT	CAAAACAT	CAGAACTC	2115
Qy	811	Gly	Leu	Lys	Pro	Phe	Thr	830
Db	2116	GGTCTAG	ACCAAC	ACCAATAC	GAATTTG	CGGTTG	CATGTG	2175
Qy	831	Gly	Pro	Phe	Gly	Ser	Val	850
Db	2176	AGTCCTT	GGAGCCCT	GTAGTCT	ACCATTCT	ACTCTTCC	AGAAGC	2235
Qy	851	Ser	Asp	Leu	Arg	Leu	Ser	870
Db	2236	GTTGGAG	TAAAGTG	ACATTAAT	AGAGGAT	GACACAT	CGCCCTG	2295
Qy	871	Thr	Glu	Pro	Asn	Gly	Glu	890
Db	2296	GATGCC	CCAGAA	CAGTGTG	CACCGCT	ATACTAT	CTTATATG	2355
Qy	891	Pro	Glu	His	Gln	Thr	Leu	910
Db	2356	ATTGCAG	GAGAGT	GGCAGGCT	CTAC	CCGTGA	AGGGCA	2415
Qy	911	Gly	Leu	Glu	Ser	Asp	Thr	930
Db	2416	AAC	TGTG	TAGCAG	GAATGTG	CAATGT	CAAGAT	2475
Qy	931	Gly	Pro	Phe	Ser	Arg	Leu	949
Db	2476	GGACCC	CTTTTCA	AATCTGT	GGAGCT	GGCAGT	ACTTCC	2535
Qy	950	-----	-----	-----	-----	-----	-----	950
Db	2536	CAGAGG	CCCCA	AGCGTTT	TGATTC	TGTCGAT	GCCAAAG	2595
Qy	951	Asp	Val	His	Ala	Val	Thr	970
Db	2596	GACCA	AAATCA	ATGACT	TGGCA	TTGCTGT	AGGTGT	2655
Qy	971	Ala	Cys	Met	Cys	Ala	Gly	985

Db	2656	ATCTGTGTTCTCACTTGGATATACCGAAGTAAGCCAGGAATAATCATCTGCTTCCAGACG	2715
Qy	986	-----LeuProGlyLeuSerSerSerGlyThrProGlyAsn	997
Db	2716	GCACAGAATGGAACTCAACAGTTACCTCGTACCAGTGCCTCTCTAGCTAGTGGAAAT	2772
RESULT 12			
AAD10023			
ID	AAD10023	standard; cDNA; 825 BP.	
AC	AAD10023;		
XX			
DT	12-SEP-2001	(first entry)	
XX			
DE	Mouse Nope	(neighbour of punc ell) intracellular domain cDNA.	
XX			
KW	Mouse; Nope;	neighbour of punc ell; cytosatic; neuroprotective; vaccine;	
KW	gene therapy;	cerebroprotective; colonic cancer; mental retardation;	
KW	tumour suppressor;	chromosome 9; transgenic animal; genetic disorder;	
KW	obesity; Bardet-Biedl syndrome;	autosomal recessive disorder; retinitis;	
KW	polydactyly; pigmentosa;	hypogonadism; hypertension; diabetes mellitus;	
KW	renal anomaly;	cardiovascular anomaly; intracellular domain; ss.	
OS	Mus musculus.		
XX			
Key	Location/Qualifiers		
FT	1..825		
FT	/*tag= a		
FT	/product= "Mouse Nope intracellular domain"		
FT	/notes= "CDS does not include start and stop codon"		
FT	/partial		
XX			
PN	WO200149714-A2.		
XX			
PD	12-JUL-2001.		
XX			
PF	26-OCT-2000; 2000WO-US29698.		
XX			
PR	04-JAN-2000; 2000US-0174496.		
PR	19-MAY-2000; 2000US-0205789.		
XX			
PA	(NEUR-) NEUROSCIENCES RES FOUND INC.		
PI	Salbaum JW;		
PI	WPI; 2001-441846/47.		
DR	P-FSDS; AAE05253.		
DR			
XX			
PT	Murine Nope polypeptides and nucleic acids useful for preventing,		
PT	diagnosing and treating colonic cancer and Bardet-Biedl syndrome -		
XX			
PS	Claim 11; Page 83-85; 99pp; English.		
XX			
CC	The present invention relates to Nope (neighbour of punc ell) which is		
CC	used in the prevention, treatment and diagnosis of diseases associated		
CC	with inappropriate Nope expression such as cancers especially colonic		
CC	cancer and genetic disorders as Nope is thought to be a tumour		
CC	suppressor. Nope gene is located on chromosome 9 and is used in gene		
CC	therapy. Nope is used as vaccine. Nope gene may be administered to treat		
CC	diseases by rectifying mutations or deletions in a patient's genome that		
CC	affect the activity of Nope by expressing inactive proteins or to		
CC	supplement the patients own production of Nope polypeptides. Nope gene		
CC	is used to study the expression and function of Nope polypeptides and		
CC	their role in metabolism through the creation of transgenic animal		
CC	models. The anti-Nope antibodies and Nope antagonists may also be used		
CC	to down regulate Nope expression and activity for the treatment of		
CC	Bardet-Biedl syndrome which is an autosomal recessive disorder		
CC	characterised by mental retardation, obesity, polydactyly, retinitis		
CC	pigmentosa and hypogonadism. Patients with Bardet-Biedl syndrome have a		
CC	high incidence of hypertension, diabetes mellitus and renal and		
CC	cardiovascular anomalies. The present sequence is mouse Nope (neighbour		
CC	of punc ell) intracellular domain cDNA.		
XX			

SQ Sequence 825 BP; 174 A; 271 C; 236 G; 144 T; 0 other;

Alignment Scores:

Pred. No.: 2,06e-62 Length: 825
Score: 1472.00 Matches: 275
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 22.29% Indels: 0
DB: 22 Gaps: 0

US-09-754-997A-2 (1-1252) x AADI0023 (1-825)

```
QY 978 ArgGlnSerSerHisArgGluAlaLeuProGlyLeuSerSerSerGlyThrProGlyAen 997
Db 1 CGCAAAAGCTCCACAGGAAGCCCTTCCCGATTGTCTCTCAGGCACCCAGGAAC 60

QY 998 ProAlaLeuTyrThrArgAlaArgLeuGlyProProSerValProAlaAlaHisGluLeu 1017
Db 61 CCAGCGCTCTACACAAGAGCTCGGCTTGGGCTCTCCAGTGCTCCCTGCTGCTCATGAGTTG 120

QY 1018 GluSerLeuValHisProArgProGlnAspTrpSerProProSerAspValGluAsp 1037
Db 121 GAGTCCCTCGTGATCTCTGTCCTCCAGGATGGTCCCAACACCTCAGATGTGAAGAC 180

QY 1038 LysAlaGluValHisSerLeuMetGlyGlySerValSerAspCysArgGlyHisSerIys 1057
Db 181 AAGGCTGAAGTACACAGCCCTTATGGTGGCAGTGTTCAGATTGCCGGGCGCACTCCCAAG 240

QY 1058 ArgIysIleSerTrpAlaGlnAlaGlyClyProAntTrpAlaGlySerTrpAlaGlyCys 1077
Db 241 AGAAGATCTCTGGGCTCAGCGAGGGGGACAACTGGGCGAGGCTCTGGGCGAGGCTGT 300

QY 1078 GluLeuProGlnGlySerGlyProArgProAlaLeuThrArgAlaLeuLeuProAla 1097
Db 301 GAGTGCCCCAGGGTAGTGTGTCACAGGCGGCTCTGACCCCTGTCTGCTGCTCCAGCG 360

QY 1098 GlyThrGlyGlnThrLeuLeuGlnAlaLeuValTyrAspGlyIleLysSerAenGly 1117
Db 361 GGAACCGGGCAGACACTCTGCTGCAAGCCCTGTGTATGACGGCATAAGAGCAAGCGG 420

QY 1118 ArgIysIleProSerProAlaCysArgAenGlnValGluAlaGluValIleValHisSer 1137
Db 421 AGAAGAGAGCGTGTCCCAAGCTGCGAGGAATCAGGTGGAAGCTGAGGTCAATGTCCACTCC 480

QY 1138 AspPheGlyAlaSerLysGlyCysProAspLeuHisLeuGlnAspLeuGluProGluGlu 1157
Db 481 GACTCGGTGATCCAAAGAGATGCTCTGACTCCACCTCCAGACCTGGAGCCAGAGGAA 540

QY 1158 ProLeuThrAlaGluThrLeuProSerThrSerGlyAlaValAspLeuSerGlnGlyAla 1177
Db 541 CCACTGACTGCAGAGACTCTGCTTCCACGCTGTGAGCTGTGGATCTGTCTCAAGGAGCA 600

QY 1178 AspTrpLeuGlyArgGluLeuGlyClyCysGlnProThrThrSerGlyProGluArgLeu 1197
Db 601 GACTGGCTGGGCGAGGAGCTGGAGGGTGGCAACCAACACAGTCAGTGGGCCAGAGAGGCTC 660

QY 1198 ThrCysLeuProGluAlaAlaSerAlaSerCysSerCysSerAspLeuGlnProSerThr 1217
Db 661 ACCTGCTTGCACAGAGCAGCAGTGTCTCTGCTCTGCTCAGACCTCCACCCAGCACT 720

QY 1218 AlaIleGluAlaProGlyLysSerCysGlnProLysAlaLeuCysProLeuThrVal 1237
Db 721 GCTATAGAGGAGGCCCTTGGGAAGAGCTGCCAGGCCCAAGCCCTGTGTCTCTTAACATC 780

QY 1238 SerProSerLeuProArgAlaProValSerSerAlaGlnValPro 1252
Db 781 AGCCCAAGCCTTCCAGGGGCCCTGTCTCTCTGTCTGCTCAGGTCCCC 825
```

RESULT 13

AA576658.

ID AA576658 standard; cDNA; 3330 BP.

XX

AC AA576658;

XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #12462.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KM food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSB-) HYSBQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR P-PSDB; ABG12471.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -

XX PS Claim 1; SEQ ID No 12462; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human
XX CC diagnostic coding sequences of the invention.

XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 3330 BP; 971 A; 775 C; 746 G; 838 T; 0 other;

Alignment Scores:

Pred. No.: 4,05e-61 Length: 3330
Score: 1458.50 Matches: 341
Percent Similarity: 50.58% Conservatives: 139
Best Local Similarity: 35.93% Mismatches: 324
Query Match: 22.09% Indels: 145
DB: 23 Gaps: 16

US-09-754-997A-2 (1-1252) x AAS76658 (1-3330)

```
QY 131 SerGlnValAlaValValLysLeuAlaThrLeuGluAspPheSerLeuHisProGluSer 150
Db 226 AGTAGCTCAATGTTAGTGAACACCACTTCTATCTCTGATTCAGTCCAGCCATTTCC 285

QY 151 GlnIleValGluGluAenGlyThrAlaArgPheGluCysHisThrIysGlyLeuProAla 170
Db 151 GlnIleValGluGluAenGlyThrAlaArgPheGluCysHisThrIysGlyLeuProAla 170
```


Qy	187	ArgLeuIleThrLeuProGlyTyrTrpLeuGlnIleLeuAspValGlnAspSerAspAla	206
Db	541	AGGATAACTGCCTTACCACAGGAGTATTGGCAGATCTATGATGTTCAGCCAAAGGGATTCT	600
Qy	207	GlySerTyrArgCysValAlaThrAenSerAlaArgGlnArgPheSerGlnGluAlaSer	226
Db	601	GGAATATATCGTTGTATTGTCTGCCACTGTAGCCACCAGCATGAAGATATGGAGGCTTCG	660
Qy	227	LeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAspValValIleVal	246
Db	661	CTAACTGTGATT-----CCAGCTAAGGAGTCAAAATCCTTCCACACACCAACCAATTATA	714
Qy	247	AlaAlaProGluAenThrThrValValSerGlyGlnAenValValMetGluCysValAla	266
Db	715	GCAGGTCACAGACATACACACATCTCTTCATCAGACTGTGTAGTTTTTGGATATGATGCC	774
Qy	267	SerAlaAepProThrProPheValSerTyrValArgGlnAspGlyLysProIleSerThr	286
Db	775	ACAGGAATCCCAACCCATCATTTCTTGGAGCGCGCTTGATCAAAATCCATTTGATGTC	834
Qy	287	--- AspValIleValLeuGlyArgThrAenLeuLeuAlaSerAlaGlnProArgHis	305
Db	835	TTTAAATACTCGGGTACTTGGAAATGGTAATCTCATGATATCTGATCAAAATCCATTTGATGTC	894
Qy	306	SerGlyValTyrValCysArgAlaAenLysProLeuThrArgAspPheAlaThrAlaAla	325
Db	895	GCTGGAGTATATGTTGTGCGGCGCACTACCCCTGGCACACCAACTTTTACAGTTGCTATG	954
Qy	326	AlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerArg	345
Db	955	GCACTTTTACTGTATTAGTCTCTCTTCATTGTTGAAATGGCCAGAAAGATTAAACAGG	1014
Qy	346	ThrArgAlaSerThrAlaAArgPheValCysArgAlaSerGlyLysProArgProAlaLeu	365
Db	1015	CCTCGAGCTGGCACTGCTCGATTGTTGTGTGTCAGCAGAGAAGAAATCCCTCTCCCAAGATG	1074
Qy	366	HisTrpLeuHisAspGlyIleProLeuArgProAenGlyArgValLysValGlnGlyCly	385
Db	1075	TCATGTTTGAATAATGGAAAGAAATACATTCGAATGGTAGAATTAATAATGTGAC-----	1128
Qy	386	GlyGlySerLeuValIleThrGlnIleGlyLeuGlnAepAlaGlyTyrTyrGlnCysVal	405
Db	1129	AACAGTAAATTTGGTAAATTAACCAAGATTATCTCTGAAGATGATGCTATTTATCAGTGGATG	1188
Qy	406	AlaGluAenSerAlaGlyThrAlaCysAlaAlaProLeuAlaValValAlaArgGlu	425
Db	1189	GCTCAGAATAGCCAAAGGATCTATTTATCTAGAGCCAGACTGACTGTAGTGTGATGTCAGAA	1248
Qy	426	GlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerSerValLeu	445
Db	1249	GACAGACCAGTGTCTCCCTAATAATGTACATGCTGAAACCAATGTCAAGCTCAGCCATTCCT	1308
Qy	446	ValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTyr	465
Db	1309	TTAGCTGGGAGAGCCACTTTTAAATTCAGACAAAGTATCTGCTTATCTGTACACTAC	1368
Qy	466	GlnLysAlaArgGlyValAspAsnValGluTyrGlnPheAlaValAenAsnAspThrThr	485
Db	1369	ATGAAGCAGAGGTTTAAATAATGAAGAGATATCAAGTAGTCACTCGGAATATGACACAACT	1428
Qy	486	GluLeuGlnValArgAspLeuGluProAenThrAspTyrGluPheTyrValValAlaTyr	505
Db	1429	CATTATATTATGATGACTTAGAGCCTGCCAGCAATATATCTTCTACATTGTAGCATAT	1488
Qy	506	SerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThrLeuAspAspVal	525
Db	1489	ATGCCAATGGGAGCCAGCCAGATGCTCTCACCATGTGACACAGATACTCTAGAGGATGAC	1548
Qy	526	ProSerAlaAlaProGlnLeuThrLeuSerSerProAenProSerAspIleArgValAla	545
Db	1549	CCCAAGA-----	1554

Db 2128 GGGCAATACCATGGCTTTGCTAGAAAACTTGGTAGCAGGAATGTACATTTGTCAAG 2187
 Qy 922 MetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIleThr 941
 Db 2188 ATATCTGCATCCAAATGAGTGGGAGAGACCCCTTTTCAAAATTCCTGGAGCTGGCAGTA 2247
 Qy 942 LeuGlnGluThrPheSerAspSer----- 949
 Db 2248 CTTCAAGGAAGAACCTTGAAATCAATCAGAGGCCCAAGCGTTTAGATTTCTGCTGATGCC 2307
 Qy 950 -----LeuAspValHisAlaValThrGlyIleIleValGly 961
 Db 2308 AAAGTTTATTACCATTTACCATCTGGACCAAAAATCAATGACTGGCATTGCTGTAGGT 2367
 Qy 962 ValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGlnSerSer 981
 Db 2368 GTTGGCATAGCTTTGACCTGCATCTCATCTGTGTCTTCTCATCTTGTATATACCGAAGTAAA 2427
 Qy 982 HisArgGluAla-----LeuProGlyLeu 989
 Db 2428 GCCAGGAATCATCTGCTTCCAGACGCCACAGATGAAGTCAACAGTTACCTCGTACC 2487
 Qy 990 SerSerSerGlyThrProGlyAsn 997
 Db 2488 AGTCCCTCTTAGCTAGTAGTGGAAAT 2511
 RESULT 15
 ABX71436
 ID ABX71436 standard; cDNA; 2958 BP.
 XX
 AC ABX71436;
 XX
 DT 06-MAR-2003 (first entry)
 XX
 DE Human cDNA encoding a novel tyrosine phosphatase, NHP5.
 KW Human; ss; gene; chromosome 15; chromosome 3; tyrosine phosphatase;
 KW Ig super family protein; gene therapy; NHP; novel human protein.
 XX Homo sapiens.
 OS
 XX US6465632-B1.
 PN
 XX 15-OCT-2002.
 PD
 XX 08-JUN-2001; 2001US-0877730.
 PF
 XX 09-JUN-2000; 2000US-210607P.
 PR
 XX (LEXI-) LEXICON GENETICS INC.
 PA
 XX Walke DW, Scoville J, Turner CA, Friedrich G, Abuin A;
 PI Zambrowicz B, Sands AT;
 PI
 XX WPI; 2003-147071/14.
 DR
 XX P-PSDB; ABUS4204.
 PT Novel isolated nucleic acid which encodes a novel human protein that
 PT shares sequence similarity with animal phosphatases, that is useful for
 PT generating antibodies, and as reagents in diagnostic assays -
 XX
 PS Disclosure; Column 41-44; 58pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule comprising a
 CC nucleotide sequence encoding a novel human protein (NHP) that shares
 CC sequence similarity with animal phosphatases (in particular tyrosine
 CC phosphatases and is a member of the Ig superfamily). The disclosed NHP
 CC polynucleotide sequences (from genes located on either chromosome 15
 CC or chromosome 3) are useful for identifying coding sequences and in the
 CC identification of biologically relevant splice junctions. The NHP
 CC polynucleotide sequences are useful in gene therapy, for detecting
 CC mutant NHPs or inappropriately expressed NHPs, for the diagnosis of

CC disease, for screening drugs effective in treatment of symptomatic or
 CC phenotypic manifestations of perturbing the normal function of NHP.
 CC Complementary sequences of the NHP polynucleotides used in conjunction
 CC with PCR to screen libraries, isolate clones and prepare cloning and
 CC sequencing templates. Labeled NHP nucleotide probes can be used to screen
 CC a human genomic library which is helpful for identifying polymorphisms,
 CC determining the genomic structure of a given locus/allele and designing
 CC diagnostic tests. The probe sequences also have use in defining and
 CC monitoring both drug action and toxicity. Oligonucleotides
 CC complementary to NHPs may encode or act as NHP antisense molecules, or
 CC may be used as part of ribozyme and/or triple helix sequences.
 CC The present sequence encodes an NHP.
 XX
 SQ Sequence 2958 BP; 855 A; 694 C; 643 G; 764 T; 2 other;
 Alignment Scores:
 Pred. No.: 1,33e-57 Length: 2958
 Score: 1384.00 Matches: 334
 Percent Similarity: 46.11% Conservative: 140
 Best Local Similarity: 32.49% Mismatches: 244
 Query Match: 20.96% Indels: 24
 DB: 25 Gaps: 14
 US-09-754-997A-2 (1-1252) x ABX71436 (1-2958)
 Qy 9 GlyLeuLeuValLeuThrPheCysLeuSerAlaArgGlyGluLeuProLeuProGln 28
 Db 43 GGGATGCTGCTCGCGCGCTCTGCTCTCTGCTGCTCAGT-----CCTTTGCCAGGA 96
 Qy 29 GluThrThrVal---LysLeuSerCysAspGluGlyProLeuGlnValIleLeuGlyPro 47
 Db 97 GTGTGGTGTCTTAGCGAACTGTCTTTGTAAGAACACACAGCATGTAACTGTACACAGA 156
 Qy 48 GluGlnAlaValValLeuAspCysThrLeuGlyAlaThrAlaAlaGlyProThrArg 67
 Db 157 AAGGACCCAGTCTGCTTTTAGATTTGCCAGGCTCACGGAGAGTT-----CCTATTAA 207
 Qy 68 ValThrTrpSerLysAspGlyAspThrValLeuGluHisGluAlaLeuHisLeuLeuPro 87
 Db 208 GTACATGGTTGAATAATGAGGCAAAAATGCTGTGAAATATAACGGATCGAGGTCTTCT 267
 Qy 88 AsnGlySerLeuTrpLeuSerSerProLeuGluGlnGluAspSerAspAspGluAla 107
 Db 268 AACGGCTCTTTATACATCAGT-----GAGTGAAGGCGGAGGCGAGGAGCAGTCC 318
 Qy 108 LeuArgIleTrpLysValThrGluGlySerTyrSerCysLeuAlaHisSerProLeuGly 127
 Db 319 -----GATGAAGGATTTTATCAGTGTCTGGCAATGAACAAATATGGA 360
 Qy 128 ValValAlaSerGlnValAlaValValLysLeuAlaThrLeuGluAspPheSerLeuHis 147
 Db 361 GCCATTCTTAGTCAAAAGCTCATCTTGCCTTATCAACTATTCTGCAATTGAGTCCAG 420
 Qy 148 ProGluSerGlnIleValGluGluAsnGlyThrAlaArgPheGluCysHisThrLysGly 167
 Db 421 CCAATTTCCACTGAGGTCCACGAAGGTGGAGTGTCTGATTTGCATTCAGATTTTCATCC 480
 Qy 168 LeuProAlaProIleIleThrTrpGluLysAspGlnValThrValPro---GluGluPro 186
 Db 481 CACCCCTCTGCAGTCATATACATGGAGTTTCAATCGGACAACTCTTACCTATGACTATGAC 540
 Qy 187 ArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuAspValGlnAspSerAspAla 206
 Db 541 AGGATACTGCCCTTACCACAGGAGTATTGAGATCTATGATGTCACCCAAAGGGATTTCT 600
 Qy 207 GlySerTyrArgCysValAlaThrAsnSerAlaArgGlnArgPheSerGlnGluAlaSer 226
 Db 601 GGAAATTCGTGTATTGCTGCCACTGTAGCCACCGACGTAAGAGTATGAGGCGCTCG 660
 Qy 227 LeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAspValIleVal 246
 Db 661 CTAATCTGTGATT-----CCAGCTAAGGAGTCAAAATCCTTCCACACACACCAATATA 714


```
Db 2308 AAAGTTTATTGAGGATATTACCATCTGGACCCAAAAATCAATGACTGGCAITTCCTGTAGGT 2367
Qy 962 ValCysLeuGlyLeuLeuCysLeuAlaCysMetCysAlaGlyLeuArgGlnSerSer 981
Db 2368 GTTGGCATAGCCTTGACCTGCATCCTCATCTGTGTCTTCATCTTTGATATACCGAAGTAA 2427
Qy 982 HisArgGluAla-----LeuProGlyLeu 989
Db 2428 GCCAGGAATCATCTGCTTCCAGACGGCACAGAAATGGAACCTCAACAGTTACCTCGTACC 2487
Qy 990 SerSerSerGlyThrProGlyAsn 997
Db 2488 AGTGCCTCCTTAGCTAGTGGAAAT 2511
```

Search completed: October 14, 2003, 06:22:47
Job time : 680 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 14, 2003, 06:11:10 ; Search time 872 Seconds

(without alignments)
3725.578 Million cell updates/sec

Title: US-09-754-997A-2

Perfect score: 6604

Sequence: 1 MARADTGRGLLVLTCLLSA.....CPLTVSPSLPRAPVSSAQVP 1252

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1731049 seqs, 1297405648 residues

Total number of hits satisfying chosen parameters: 3462098

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09754997/runat_14102003_061101_9762/app_query.fasta_1.1415
-DB=Published Applications NA -QWFI=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODS=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09754997@cgn_1_1_243@runat_14102003_061101_9762
-NCPUS=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6604	100.0	3756	10	US-09-754-997A-45 Sequence 45, Appl

2	6604	100.0	6176	10	US-09-754-997A-1	Sequence 1, Appl
3	5746	87.0	3753	12	US-10-239-663-29	Sequence 29, Appl
4	5610	84.9	3741	10	US-09-908-193-1	Sequence 1, Appl
5	4903	74.2	2796	10	US-09-754-997A-3	Sequence 3, Appl
6	3167.5	48.0	2223	12	US-10-239-663-28	Sequence 28, Appl
7	1747.5	26.5	3453	14	US-10-231-353-7	Sequence 7, Appl
8	1746.5	26.4	3210	14	US-10-231-353-1	Sequence 1, Appl
9	1746.5	26.4	3874	14	US-10-231-353-31	Sequence 31, Appl
10	1691.5	25.6	2976	14	US-10-231-353-11	Sequence 11, Appl
11	1691.5	25.6	3219	14	US-10-231-353-17	Sequence 17, Appl
12	1472	22.3	825	10	US-09-754-997A-5	Sequence 5, Appl
13	1384	21.0	2715	14	US-10-231-353-5	Sequence 5, Appl
14	1384	21.0	2958	14	US-10-231-353-9	Sequence 9, Appl
15	1329	20.1	2481	14	US-10-231-353-15	Sequence 15, Appl
16	1329	20.1	2724	14	US-10-231-353-19	Sequence 19, Appl
17	1144	17.3	2139	14	US-10-231-353-21	Sequence 21, Appl
18	1144	17.3	2382	14	US-10-231-353-27	Sequence 27, Appl
19	951.5	14.4	5297	12	US-10-007-926A-270	Sequence 270, App
20	909	13.8	1875	14	US-10-231-353-23	Sequence 23, Appl
21	781.5	11.8	1644	14	US-10-231-353-25	Sequence 25, Appl
22	781.5	11.8	1887	14	US-10-231-353-29	Sequence 29, Appl
23	647.5	9.8	1143	14	US-10-231-353-3	Sequence 3, Appl
24	624.5	9.5	7718	14	US-10-175-523-192	Sequence 192, App
25	607	9.2	8232	14	US-10-198-846-10976	Sequence 10976, A
26	592.5	9.0	909	14	US-10-231-353-13	Sequence 13, Appl
27	584.5	8.9	4277	11	US-09-946-374-57	Sequence 57, Appl
28	584.5	8.9	4277	12	US-10-015-387A-57	Sequence 57, Appl
29	584.5	8.9	4277	12	US-10-006-130A-57	Sequence 57, Appl
30	584.5	8.9	4277	12	US-10-199-672-439	Sequence 439, App
31	584.5	8.9	4277	12	US-10-006-172A-57	Sequence 57, Appl
32	584.5	8.9	4277	12	US-10-187-749-439	Sequence 439, App
33	584.5	8.9	4277	12	US-10-194-457-439	Sequence 439, App
34	584.5	8.9	4277	12	US-10-184-642-439	Sequence 439, App
35	584.5	8.9	4277	12	US-10-196-747-439	Sequence 439, App
36	584.5	8.9	4277	12	US-10-015-382A-57	Sequence 57, Appl
37	584.5	8.9	4277	12	US-10-017-253A-57	Sequence 57, Appl
38	584.5	8.9	4277	12	US-10-173-689-439	Sequence 439, App
39	584.5	8.9	4277	12	US-10-173-690-439	Sequence 439, App
40	584.5	8.9	4277	12	US-10-173-691-439	Sequence 439, App
41	584.5	8.9	4277	12	US-10-173-692-439	Sequence 439, App
42	584.5	8.9	4277	12	US-10-173-694-439	Sequence 439, App
43	584.5	8.9	4277	12	US-10-173-698-439	Sequence 439, App
44	584.5	8.9	4277	12	US-10-173-699-439	Sequence 439, App
45	584.5	8.9	4277	12	US-10-173-707-439	Sequence 439, App

ALIGNMENTS

RESULT 1

US-09-754-997A-45
; Sequence 45, Application US/09754997A
; Patent No. US20020102551A1
; GENERAL INFORMATION:
; APPLICANT: Salbaum, Michael J.
; TITLE OF INVENTION: NOPE Polypeptides, Encoding Nucleic
; TITLE OF INVENTION: Acids and Methods of Use
; FILE REFERENCE: P-NI 4552
; CURRENT APPLICATION NUMBER: US/09/754,997A
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: US 60/174,496
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: US 60/205,789
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 3756
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-754-997A-45

Alignment Scores:

Pred. No.: 0 Length: 3756

Db 2041 GGTCTGGAGATCAAGCTTGGGACGCTGGGGCCGCTGGAGTGAAGCAG 2100
Qy TyrGluLeuThrGlnLeuValProGlyArgProTyrGluValLysLeuValAlaPheAen 720
Db 2101 TATGAATCAGCCAGTATGCTTCCCTGGCAGGCGGTACAGAGTGAAGCTTCTCAAC 2160
Qy LysHisGluAspGlyTyrAlaValTyrLysGlyLysThrGluLysAlaProThrPro 740
Db 2161 AAACAGAGGAGCGGTACCTCTCTGTGTGGAAGGGCAACAGGAGAGGCGCCACGCA 2220
Qy AspLeuProIleGlnArgGlyProProLeuProProAlaHisValHisAlaGluSerAen 760
Db 2221 GACCTGCTATCCAGAGGGGGCCACCGCTCTCTGCCCCATGTCCACGACAGTCAAC 2280
Qy SerSerThrSerIleTyrLeuArgTyrLysLysProAspPheThrThrValLysIleVal 780
Db 2281 AGCTCCACTTCCATTTGGTCTTGGTGAAGAGCCAGACTTTACCACTGTCAAGATTGTC 2340
Qy 781 AsnTyrThrValArgPheGlyProTyrGlyLeuArgAsnAlaSerLeuValThrTyr 800
Db 2341 AACTACACTGTACGCTTCCGCCCCCTGGGGGCTCAGGAATGCTTCCCTGGTCACTACTAT 2400
Qy 801 ThrSerSerGlyGluAspIleLeuIleGlyLysLeuLysProPheThrLysTyrGluPhe 820
Db 2401 ACCAGCTCTGGAGAGACATTTCTCATTTGGCGGCTGAAACCAATTTACCAAGTACAGATT 2460
Qy 821 AlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSer 840
Db 2461 GCGGTACAGTCCACGAGTGGATATGATGGGCTTTGGCTCCGTGTAGAACGCTCC 2520
Qy 841 ThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrPro 860
Db 2521 ACCCTGCTTACCGGCTTTCAACACCTCTCTGACCTCGGCTCGGCTCGGCTCGGCTCG 2580
Qy 861 SerThrValArgLeuHisTyrCysProThrGluProAsnGlyGluIleValGluTyr 880
Db 2581 TCACCGTTTGGTTCACCTTGGTGTCCCCCAGAGCCCAATGGTGAGATTGTGGAGTAT 2640
Qy 881 LeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTyrThrLeuLeuThrThr 900
Db 2641 CTAATTTCTCTACAGCAACCAACACACCCAGCCCGAACACCACTGGTCAACCA 2700
Qy 901 GluGlyAsnIlePheSerIleGluValHisGlyLeuGluSerAspThrArgTyrPhePhe 920
Db 2701 GAGGGAAACATCTTCAGTGCAGAGGTCCATGGCTAGAGAGTCAACCTGGTATTCTTC 2760
Qy 921 LysMetGlyValArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIle 940
Db 2761 AAGATGGAGCCCGCACAGGTGGGCGCTGGGCGCTTTTCCGCTTGCAGGATGTGATT 2820
Qy 941 ThrLeuGlnGluThrPheSerAspSerLeuAspValHisAlaValThrGlyIleIleVal 960
Db 2821 ACTCTGCAAGAGACATTTCTCAGACTCTTGGATGTGCACGCGTCAOGGCGCATCTG 2880
Qy 961 GlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGlnSer 980
Db 2881 GGTGTCTGCTGGGCTTCTCTGCTCTCTGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 2940
Qy 981 SerHisArgGluAlaLeuProGlyLeuSerSerSerGlyThrProGlyAsnProAlaLeu 1000
Db 2941 TCCACAGGAGAGCCCTTCCCGATTGTCTCTCAGGACCCCGAACCAACCCAGCGCTC 3000
Qy 1001 TyrThrArgAlaArgLeuGlyProProSerValProAlaAlaHisGluLeuGluSerLeu 1020
Db 3001 TACACAGAGCTCGGCTTGGGCTCCAGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 3060
Qy 1021 ValHisProArgProGlnAspTyrSerProProProSerAspValGluAspLysAlaGlu 1040
Db 3061 GTGCTATCTCTGCTCCCGAGATTGGTCCCCACCACTCAGATGTGCAAGCAAGGCTGAA 3120
Qy 1041 ValHisSerLeuMetClyGlySerValSerAspCysArgGlyHisSerLysArgLysIle 1060
Db 3121 GTACACAGCTTATGGTGGCAGTGTTCAGATTGCCGGGGCCACTTCCAAAGAGAAGATC 3180

Qy 1061 SerTrpAlaGlnAlaGlyGlyProAsnTrpAlaGlySerTrpAlaGlyCysGlnLeuPro 1080
Db 3181 TCCTGGGCTCAGGCAGGGGACCAAACTGGGCGAGGCTCTCTGGCAGGCTGTGAGCTGCC 3240
Qy 1081 GlnGlySerGlyProAspProAlaLeuThrArgAlaLeuLeuProProAlaGlyThrGly 1100
Db 3241 CAGGATGTGTGGTCCAGGCGCGCTCTGACCGTGTCTGTGCTCTCCAGCGGAAACCGGG 3300
Qy 1101 GlnThrLeuLeuLeuGlnAlaLeuValTyrAspGlyIleLysSerAsnGlyArgLysLys 1120
Db 3301 CAGACACTGCTGCTGCAAGCCCTGTGTATGACGCGCATAAAGAGCAACCGGAGAAAGAG 3360
Qy 1121 ProSerProAlaCysArgAsnGlnValGluAlaGluValIleValHisSerAspPheGly 1140
Db 3361 CGTCCCCAGGCTGCGAGGAATCAGGTGGAAGCTGAGGTCAATTGTCTCACTCCGACTTCGGT 3420
Qy 1141 AlaSerLysGlyCysProAspLeuHisLeuGlnAspLeuGluProGluGluProLeuThr 1160
Db 3421 GCATCCAAAGATGTCTGACCTCCACCTCAAGACCTGGAGCCAGAGAAACCACTGACT 3480
Qy 1161 AlaGluThrLeuProSerThrSerGlyAlaValAspLeuSerGlnGlyAlaAspTrpLeu 1180
Db 3481 GCAGAGACTCTGCTCTCCAGCTCTGGAGCTGTGGATCTGTCTCAAGGAGCAGACTGGCTG 3540
Qy 1181 GlyArgGluLeuGlyGlyCysGlnProThrThrSerGlyProGluArgLeuThrCysLeu 1200
Db 3541 GGCAGGAGCTGGGAGGCTGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 3600
Qy 1201 ProGluAlaAlaSerAlaSerCysSerCysSerAspLeuGlnProSerThrAlaIleGlu 1220
Db 3601 CCAGAGCAGCAGCAGTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 3660
Qy 1221 GluAlaProGlyLysSerCysGlnProLysAlaLeuCysProLeuThrValSerProSer 1240
Db 3661 GAGGCGCTTGGAAAGTGTGCGCCCAAGCCCTGTGTCTCTTAAAGTCAAGCCCAAGC 3720

RESULT 2

US-09-754-997A-1
; Sequence 1, Application US/09754997A
; Patent No. US20020102551A1
; GENERAL INFORMATION:
; APPLICANT: Salbaum, Michael J.
; TITLE OF INVENTION: NOPE Polypeptides, Encoding Nucleic
; TITLE OF INVENTION: Acids and Methods of Use
; FILE REFERENCE: P-NI 4552
; CURRENT APPLICATION NUMBER: US/09/754,997A
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: US 60/174,496
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: US 60/205,789
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 6176
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (3756)
US-09-754-997A-1
Alignment Scores:
Pred. No.: 0 Length: 6176
Score: 6604.00 Matches: 1252
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

DB:	10	Gaps:	0
US-09-754-997A-2 (1-1252) x US-09-754-997A-1 (1-6176)			
Qy	1	MetAlaArgAlaAspThrGlyArgGlyLeuLeuValLeuThrPheCysLeuLeuSerAla	20
Db	1	ATGGCGGGCGGACACGGCGCGGGCTCTGGTGTGACCTTTCTGCTGCTGTTCGGCG	60
Qy	21	ArgGlyGluLeuProLeuProGlnGluThrThrVallyLeuSerCysAspGluGlyPro	40
Db	61	CGGGGAGAGTGCATTGGCCCGAGGACAACTGTCAAGCTGAGCTGTGATGAGGGACCC	120
Qy	41	LeuGlnValIleLeuGlyProGluGlnAlaValIleValLeuAspCysThrLeuGlyAlaThr	60
Db	121	CTGCAAGTATCTTGGGCCCTGAGCAGCTGTGGTGTGACTGCCACTTTGGGGGCTACA	180
Qy	61	AlaAlaGlyProProThrArgValThrTrpSerIysAspGlyAspThrValLeuGluHis	80
Db	181	GCTGCTGGGCTCCGACAGGGTGCATGGAGCAAGGATGGAGACACTGTACTAGAGCAT	240
Qy	81	GluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGluGlnGlu	100
Db	241	GAGAACCTGCACCTGCTACCCAAATGGCTCCCTGGCTGTCTCACCCTTAGAGCAAGAA	300
Qy	101	AspSerAspAspGluGluAlaLeuArgIleTrpIysValThrGluGlySerTrpSerCys	120
Db	301	GACAGCGATGATGAGGAGCTCTTAGATCTGGAGGCTCACTGAGGGCAGCTATTCTGT	360
Qy	121	LeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValValIysLeuAlaThr	140
Db	361	CTGGCCACACAGCCGCTAGGAGTGGTGGCCAGCCAGGTTGGTGTGCAAGCTTGCCACA	420
Qy	141	LeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThrAlaArg	160
Db	421	CTCGAAGACTTCTCTGCAACCCCGAGTCCAGATTGTGGAGGAGAACGGGACAGCAGC	480
Qy	161	PheGluCysHisThrIysGlyLeuProAlaProIleIleThrTrpGluIysAspGlnVal	180
Db	481	TTTGAAATGCCACACCAAGGGCTTCCAGCCCCCATCATATTCTGGGAAAGGACACGGT	540
Qy	181	ThrValProGluGluProArgLeuIleThrLeuProIysTrpLeuLeuGlnIleLeuAsp	200
Db	541	ACCGTGCCTGAGGAGCCCGGCTCATCACTTCCCAAGTGGCTCTCCAGATCCTAGAT	600
Qy	201	ValGlnAspSerAspAlaGlySerTrpArgCysValAlaThrAsnSerAlaArgGlnArg	220
Db	601	GTCCAGGACAGTATGACGGCTCTACCGCTGCGTGGCCACCAATTGAGCCCGCAACGA	660
Qy	221	PheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGly	240
Db	661	TTGAGCCAGGAGGCTCTGCTCACTGTGGCCCTCAGAGGCTCTTTGGAGGCTTACCAGGGG	720
Qy	241	GlnAspValValIleValAlaAlaProGluAsnThrThrValValSerGlyGlnAsnVal	260
Db	721	CAGGATGTGGTCAATTTGTGGCAGCCCGAGAGAACACCCGCTAGTGTCTGGACAGATGA	780
Qy	261	ValMetGluCysValAlaSerAlaAspProThrProPheValSerTrpValArgGlnAsp	280
Db	781	GTGATGGAGTGGTGGCTCTGCTGACCCCAACCCCTTTGTGTCTGGTCCCGCAGGAT	840
Qy	281	GlyIysProIleSerThrAspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSer	300
Db	841	GGAAGCCCTATCTCCAGGATGTATGCTCTGGGCCGAGCAATCTACTACTATGCCGAGC	900
Qy	301	AlaGlnProArgHisSerGlyValTrpValCysArgAlaAsnIysProLeuThrArgAsp	320
Db	901	GGCAGCCTCGGCATCTCTGGAGTCTATGCTCTGGCCGAGCAACAGCCCTCAGCGGTGAC	960
Qy	321	PheAlaThrAlaAlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaPro	340
Db	961	TTCCGCCATCGGGCTGCTGAGCTCCAGTGTGCTGCCCCAGCCATCTCCGAGGACCC	1020
Qy	341	GluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGlu	360
Db	1021	GAGCGCTCTCGCGACCGCGCGCGACACCGCGCGCTTCTGTGTGCGCGCGCTCCGCGAG	1080
Qy	361	ProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgVal	380
Db	1081	CCACGGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1140
Qy	381	LysValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGly	400
Db	1141	AGGTGTCAGGGCGTGGCGCAGCTTGGTCACTCAGTCCGCTCGCAGGACGCTGGC	1200
Qy	401	TrpTrpGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaAlaProLeuAla	420
Db	1201	TACTACAGTGGCTAGCAGAAAACAGCGGGGAACTGCTGTGCGCTGCGCCCTGGCG	1260
Qy	421	ValValValArgGluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSer	440
Db	1261	GTAGTGGTGGCGAGGGGCTGCCAGGCCCGCTGCTGGGTCAACGCGCCGCTGAGC	1320
Qy	441	SerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGly	460
Db	1321	AGCTCTCTGTGCTGGCTGGGAGCGGCTGAGTTGCACAGCGCAAAATCATTTGGC	1380
Qy	461	PheSerLeuHisTrpGlnIysAlaArgGlyValAspAsnValGluTrpGlnPheAlaVal	480
Db	1381	TTCTCTCTTCACTACCAAAAGGCAAGGGAGTGGCAATGTGTGAGTACCAGTTTGCAGTA	1440
Qy	481	AsnAspAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAspTrpGluPhe	500
Db	1441	AACAATGACACACAGCTGCGAGTTGGGACCTGGAAACCCCAACGGAATATAGTTTC	1500
Qy	501	TrpValValAlaTrpSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHis	520
Db	1501	TAGTGGTGGCTACTCTCCAGCTGGGGGCCAGCGGAACCTCCAGCCCGAGCCCTGGTGCAT	1560
Qy	521	ThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSer	540
Db	1561	ACACTGGACGATGCTCCCGAGCGACGCCACCTTGTGTGTGTCAGCCCAACCCCTCG	1620
Qy	541	AspIleArgValAlaTrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuIys	560
Db	1621	GACATCAGGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1680
Qy	561	TrpIysIleGlyTrpGlyLeuGlyIysGluAspGlnValPheSerThrGluValProGly	580
Db	1681	TACAGATAGTACGGTTTGGGAGGAGATGATGAGTTTCTCCAGGAGTGGCTTGA	1740
Qy	581	AsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnIysValTrpArgValArgIle	600
Db	1741	AATGAGACACACTTACGTTAAACTCACTTCAAGCAAAACAAAGTGTACCGAGTCCGAT	1800
Qy	601	SerAlaGlyThrGlyAlaGlyTrpGlyValProSerGlnTrpMetGlnHisArgThrPro	620
Db	1801	TCAGCTGGCATCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1860
Qy	621	GlyValHisAsnGlnSerHisValProPheAlaProAlaGluLeuIysValArgAlaIys	640
Db	1861	GGTGTGCACAAACAGCCCATGTTCCCTTTGCCCTCGCAGAAATGAGGTGAGGCGCAAG	1920
Qy	641	MetGluSerLeuValValSerTrpGlnProProHisProThrGlnIleSerGlyTrp	660
Db	1921	ATGGAGTCCCTGGTGGTGTCTATGGCAGCCGCCCTCACCCACCCAGATCTCTGGATAC	1980
Qy	661	LysLeuTrpTrpGlyGluValGlyThrGluGluGluAlaAspGlyAspArgProProGly	680
Db	1981	AACTCTACTGGGGAGAGTGGGAACAGAGGAGGAGGAGATGCTGTGTGACCGCCCGG	2040
Qy	681	GlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuIysValIysValIysGln	700
Db	2041	GCTGTGGAGATCAAGCTTGGAGCTCGGGCCGCTGCGCTGGAAGAGAAAGTGAAGCAG	2100
Qy	701	TyrGluLeuThrGlnLeuValProGlyArgProTrpTrpGluValIysLeuValAlaPheAsn	720

Db 2101 TATGAACGACCGAGTTCCTCGGAGCGCGTACGAGGTGAAGCTCGTAGCTTCAAC 2160
QY 721 LysHisGluAspGlyTyrAlaAlaValTrpLysGlyLysThrGluLysAlaProThrPro 740
Db 2161 AAACACGAGCGGCTACGCTGCTGTGTGAGAGGCAAGACGAGAGGCGCCACGCCA 2220
QY 741 AspLeuProIleGlnArgGlyProProLeuProProAlaHisValHisAlaGluSerAsn 760
Db 2221 GACCTCCCTATCCAGAGGGGGCCACCGCTGCCCTCTCCGCCATGTCCACGAGAGTCAAA 2280
QY 761 SerSerThrSerIleTrpLeuArgTrpLysLysProAspPheThrThrValLysIleVal 780
Db 2281 AGCTCCACTTCCATTTGGCTTGGTGGAGAGCCAGACTTTACCACTGTGCAGATTGTC 2340
QY 781 AsnTyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTyrTyr 800
Db 2341 AACTACACTGTACGCTTGGCGCCCTGGGGGCTCAGGAATGCTTCCCTGGTCACTACTAT 2400
QY 801 ThrSerSerGlyGluAspIleLeuIleGlyLysLysProPheThrLysTrpGluPhe 820
Db 2401 ACCAGCTCTGGAGAGACATTTCTATTTGGCGGCTGAAACCAATTTACCAAGTACGAGTTT 2460
QY 821 AlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSer 840
Db 2461 GCGGTAAGTCCACGAGTGGATATGGATGGGCGCTTTGGCTCCGTGTAGAACGCTCC 2520
QY 841 ThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrPro 860
Db 2521 ACCCTGCTGACCGGCTTCAACACACTCTCTCTGACCTGCGCTGAGCCCTGCACACCA 2580
QY 861 SerThrValArgLeuHisTrpCysProProThrGluProAsnGlyGluIleValGluTyr 880
Db 2581 TCCACCGTTTGGTTACACTGCTGTCCCGCCACGAGGCCCAATGGTGGAGATTGGAGTAT 2640
QY 881 LeuIleLeuTyrSerAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThrThr 900
Db 2641 CTAAATCTCTACAGCAACCAACCAACCCAGCCCGAACCAGTGGACACTGTCTCAACCA 2700
QY 901 GluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePhe 920
Db 2701 GAGGGAACATCTTCAGTGGAGAGGTCCATGGCTAGAGGTGACACTCGGTATTCTTC 2760
QY 921 LysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIle 940
Db 2761 AAGATGGAGCGCGCACAGAGTGGGCGCTGGGCGCTTTTCCGCTTGCAGAGTGTGATT 2820
QY 941 ThrLeuGlnThrPheSerAspSerLeuAspValHisAlaValThrGlyIleIleVal 960
Db 2821 ACTCTGCAAGAGACATTTCTCAGACTCTTGGATGTGCAGCGCTCAGCGGCATCATCGTG 2880
QY 961 GlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGlnSer 980
Db 2881 GGTGTCTGCTGGGCTTCTCTGCTCTGCGCTGTGATGTGCTGCTGCTGCTGCTGCTGCT 2940
QY 981 SerHisArgGluAlaLeuProGlyLeuSerSerSerGlyThrProGlyAsnProAlaLeu 1000
Db 2941 TCCACAGGGAAGCCCTTCCCGATTGTCTCTCTCAGGACACCCAGGAAACCCAGCGCTC 3000
QY 1001 TyrThrArgAlaArgLeuGlyProProSerValProAlaHisGluLeuGluSerLeu 1020
Db 3001 TACACAGAGCTCGGCTTGGCGCTTCCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3060
QY 1021 ValHisProArgProGlnAspTrpSerProProSerAspValGluAspLysAlaGlu 1040
Db 3061 GTGCATCTCTGCTCCAGGATTTGGTCCCAACACCTCTCAGATGTGGAAGCAAGGCTGAA 3120
QY 1041 ValHisSerLeuMetGlyGlySerValSerAspCysArgGlyHisSerLysArgLysIle 1060
Db 3121 GTACACAGCTTATGGTGGCAGTGTTCAGATTGGCGGGGCCACTTCCAGAGAGAGATC 3180
QY 1061 SerTrpAlaGlnAlaGlyGlyProAsnTrpAlaGlySerTrpAlaGlyCysGluLeuPro 1080
Db 3181 TCCTGGGCTCAGGACGAGGAGCAAACTGGGAGGCTCTCTGGGAGGCTGTGAGCTGCCCC 3240

QY 1081 GlnGlySerGlyProArgProAlaLeuThrArgAlaLeuLeuProProAlaGlyThrGly 1100
Db 3241 CAGGGTAGTGGTCCAAGCGCGCTCTGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300
QY 1101 GlnThrLeuLeuGlnAlaLeuValTyrAspGlyIleLysSerAsnGlyArgLysIle 1120
Db 3301 CAGACACTGCTCTGCAAGCCCTGTGTATGACGGCATAAAGAGCAACGGGAGAAAG 3360
QY 1121 ProSerProAlaCysArgAsnGlnValGluAlaGluValIleValHisSerAspPheGly 1140
Db 3361 CCGTCCCGAGCTGCGAGGATCAGGTGGAGCTGAGGTCAATGTCCACTCCGACTTCGGT 3420
QY 1141 AlaSerLysGlyCysProAspLeuHisLeuGlnAspLeuGluProGluGluProLeuThr 1160
Db 3421 GCATCCAAAGGATGTCTGACCTCCACCTCCAAAGACCTGGAGCCAGAGGACCACTGACT 3480
QY 1161 AlaGluThrLeuProSerThrSerGlyAlaValAspLeuSerGlnGlyAlaAspTrpLeu 1180
Db 3481 CGAGAGACTCTGCTTCCACGCTGTGAGCTGTGGATCTGTCTCAAGGAGCAGACTGGCTG 3540
QY 1181 GlyArgGluLeuGlyCysGlnProThrThrSerGlyProGluArgLeuThrCysLeu 1200
Db 3541 GSCAGGAGCTGGAGGGTGGCCACCAACACACAGTGGGCGCAGAGGGCTCACCTGCTTG 3600
QY 1201 ProGluAlaAlaSerAlaSerCysSerCysAspLeuGlnProSerThrAlaIleGlu 1220
Db 3601 CCAGAGCAGCCAGTGCCTCTGCTCTGCTCAGACCTCCAGCCAGCAGCTGCTATAGAG 3660
QY 1221 GluAlaProGlyLysSerCysGlnProLysAlaLeuCysProLeuThrValSerProSer 1240
Db 3661 GAGGCGCTTGGAAAGCTGGCAGGCCAAAGCCCTGTGTCTTAACTGAGTCAAGCCAGCC 3720
QY 1241 LeuProArgAlaProValSerSerAlaGlnValPro 1252
Db 3721 CTTCCAGGCGCCCTGTCTCTCTGCTCAGGTCCCC 3756

RESULT 3

US-10-239-663-29
; Sequence 29, Application US/10239663
; Publication No. US20030139572A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia, K.
; APPLICANT: Smith, Randall, F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kabinick, Karen
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50018
; CURRENT APPLICATION NUMBER: US/10/239,663
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09226
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/192,158
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,668
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/200,166
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 3753
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-239-663-29

Alignment Scores:
Pred. No.: 0 Length: 3753
Score: 5746.00 Matches: 1098
Percent Similarity: 92.09% Conservative: 54
Best Local Similarity: 87.77% Mismatches: 96

Db 2098 AGACGATATGAGTGTACCCAGCTAGTCCCTGGCCGCTGTACGAGTGAAGCTCGTGGCT 2157
Qy PheAsnLysHisGluAspGlyTyrAlaAlaValTyrLysGlyLysThrGluLysAlaPro 738
Db 2158 TTCAACAAACATGAGATGGCTATGACGAGTGTGGAGGGCAGAGCGAGAGGGCCCG 2217
Qy ThrProAspLeuProIleGlnArgGlyProProLeuProProAlaHisValHisAlaGlu 758
Db 2218 GCACACAGATGCTATCCAGAGGGGACCAACCCCTGCCCTCCAGCCCAACGCTCCATGGGAA 2277
Qy SerAsnSerSerThrSerIleTyrLeuArgTyrLysProAspPheThrThrValLys 778
Db 2278 TCAACACAGCTCCACATCCATCTGGCTTCGGTGGTGGAAAGCCAGATTTCCACAGTCAAG 2337
Qy IleValAsnTyrThrValArgPheGlyProTyrGlyLeuArgAsnAlaSerLeuValThr 798
Db 2338 ATTGTCACTACACTGTGGCTTCAGCCCTCGGGGCTCAGGATGCCCTCCCTGGTCAAC 2397
Qy TyrTyrThrSerSerGlyGluAspIleLeuIleGlyLeuLysProPheThrLysTyr 818
Db 2398 TATTACACAGTTCCTGGAGAGACATCCTCATTTGGCGGCTTGAAGCCATTCACCAATAC 2457
Qy GluPheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGlu 838
Db 2458 GAGTTTGCAGTGCAGTCTCACGGCGTGGACATGGATGGGCTTTCGGCTCTGTGGTGGAG 2517
Qy ArgSerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeu 858
Db 2518 CGCTCCACCTCGCTGACCGGCCCTCCACACCCCATCCGACTGGACTGAGGCCCCCTG 2577
Qy ThrProSerThrValArgLeuHisTyrCysProProThrGluProAsnGlyGluIleVal 878
Db 2578 ACACCGTCCACGTTTCGGCTGCACCTGTGCCCCCCACACAGAGCCCAACGGGAGATCGTG 2637
Qy GluTyrLeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTyrThrLeuLeu 898
Db 2638 GAGTATCTGATCTGTATACAGCAGCAACACACAGCAGCTGAGCACCCAGTGGACCTTGTCTC 2697
Qy ThrThrGluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyr 918
Db 2698 ACCACGAGGAAACATCTTCAGTGTCTGAGGTCCATGGCTGTGAGAGCGACATCTGGTAC 2757
Qy PhePheLysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAsp 938
Db 2758 TTCTTCAAGATGGGGCGGCACAGAGTGGGACCTGGGCTTCTCCCGCTGCAGAT 2817
Qy ValIleThrLeuGlnGluThrPheSerAspSerLeuAspValHisAlaValThrGlyIle 958
Db 2818 GTGATCACGCTCCAGGAGAAGCTGTACAGACTCGCTGGACATGCACCTCAGTCACGGGCATC 2877
Qy IleValGlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArg 978
Db 2878 ATCTGGGGTGTGTGCTGGGCTCTCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2937
Qy GlnSerSerHisArgGluAlaLeuProGlyLeuSerSerSerGlyThrProGlyAsnPro 998
Db 2938 CGCAGCCCCCAGGAGATCTCTCCAGGCTGTCTCCAGCCCTGCTCCAGCCCGGGGATCCC 2997
Qy AlaLeuTyrThrArgAlaArgLeuGlyProProSerValProAlaAlaHisGluLeuGlu 1018
Db 2998 GCGCTGTACTCCAGAGCTCGCTTGGCCCCCCCCCAGCCCCCAGCTGCCATGAATTGGAG 3057
Qy SerLeuValHisProArgProGlnAspTyrSerProProProSerAspValGluAspLys 1038
Db 3058 TCCCTTGTGCACCCCATCCAGGACTGTGCTCCCGCCACCTCCAGACGCTGGAGGACAGG 3117
Qy AlaGluValHisSerLeuMetGlyGlySerValSerAspCysArgGlyHisSerLysArg 1058
Db 3118 GCTGAGTGCACAGCCTTATGGTGGCGGTGTTCGAAAGCGCGAGTCACTCCAAAAGA 3177
Qy LysIleSerTrpAlaGlnAlaGlyGlyProAsnTrpAlaGlySerTrpAlaGlyCysGlu 1078

Db 3178 AAGATCTCTGGGCTCAACCAAGCGGGCTGAGCTGGGCTGGTTCCTGGGAGGCTGTGTAG 3237
Qy LeuProGlnGlySerGlyProArgProAlaLeuThrArgAlaLeuLeuProProAlaGly 1098
Db 3238 CTGCCCCAG--GCAGGCCCCCGCGGCTGTGACCCCGGGCTCTGCTGCCCTCTGCTGA 3294
Qy ThrGlyGlnThrLeuLeuGlnAlaLeuValTyrAspGlyIleLysSerAsnGlyArg 1118
Db 3295 ACTGGGCAGACGCTCTGTCTGCGAGGCTCTGGTGTACGACGCCATTAAGGGCAATGGAGG 3354
Qy LysLysProSerProAlaCysArgAsnGlnValGluAlaGluValIleValHisSerAsp 1138
Db 3355 AAGAAGTCAACCCCGAGCTTCAGGAAACAGGTGGAGGCTGAAGTCAATGTCTCACTCTGAC 3414
Qy PheGlyAlaSerLysGlyCysProAspLeuHisLeuGlnAspLeuGluProGluGluPro 1158
Db 3415 TTATGTGATCTTAACGGGACCCCTGACCTCCATCTCCAGACCTCGAGGCTGGAGGACCCC 3474
Qy LeuThrAlaGluThrLeuProSerThrSerGlyAlaValAspLeuSerGlnGlyAlaAsp 1178
Db 3475 CTGCTCCAGAGGCTCTCATCTCATCTCGGGTGTGGGATCCAGGGCAGGGGCGAGCC 3534
Qy TrpLeuGlyArgGluLeuGlyCysGlnProThrThrSerGlyProGluArgLeuThr 1198
Db 3535 TGGCTGGACAGGAGTGGAGGGTGTGAGCTGGCAGCCCCCGGCCAGACAGACTTACC 3594
Qy CysLeuProGluAlaAlaSerAlaSerCysSerCysSerAspLeuGlnProSerThrAla 1218
Db 3595 TGCTGCCAGGAGCCCTGGAGATAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3654
Qy IleGluGluAlaProGlyLysSerCysGlnProLysAlaLeuCysProLeuThrValSer 1238
Db 3655 CTAGAGGAGACCCCTGGAGATAGTCTGAGCTCAAAATCCCTCTCTCTCTCTCTCTCTCTCT 3714
Qy ProSerLeuProArgAlaProValSerSerAla 1249
Db 3715 CCAGCCTGCCAGATCCCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3747

RESULT 4

US-09-908-193-1
; Sequence 1, Application US/09908193
; Publication No. US20020192748A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; APPLICANT: SHIMKETS, RICHARD A.
; APPLICANT: ZERHUSEN, BRYAN
; APPLICANT: MALYANKAR, URIEL M.
; APPLICANT: PADIGARU, MURALIDHARA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: 21402-062
; CURRENT APPLICATION NUMBER: US/09/908,193
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: 60/220,273
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/221,650
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/221,233
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 60/220,912
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/218,875
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,870
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,901
; PRIOR FILING DATE: 2000-07-18
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3741
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:

OTHER INFORMATION: Description of Unknown Organism: NOVX nucleic acid
OTHER INFORMATION: with homology to NOPG-like protein
NAME/KEY: modified_base
LOCATION: (375)
OTHER INFORMATION: a, c, t, g, other or unknown
US-09-908-193-1

Alignment Scores:

Pred. No.:	0	Length:	3741
Score:	5610.00	Matches:	1086
Percent Similarity:	91.07%	Conservative:	56
Best Local Similarity:	86.60%	Mismatches:	101
Query Match:	84.95%	Indels:	13
DB:	10	Gaps:	8

US-09-754-997a-2 (1-1252) x US-09-908-193-1 (1-3741)

Qy	1	MetAlaArgAlaAspThr-----GlyArgGlyLeuLeuValLeuThrPheCysLeuLeu	18
Db	1	ATGGCGGGGGGACGCGCGCGCGCGCGCGCTCTCGCGTTGACCTTCTGCTGCTGTG	60
Qy	19	SerAlaArgGlyGluLeuProLeuProGlnGluThrThrValLysLeuSerCysAspGlu	38
Db	61	GCCGCGCGGGGAGCTGCTGTTGCCCGAGAGACGACTGTGGAGCTGAGCTGGAGTG	120
Qy	39	GlyProLeuGlnValLeuLeuGlyProGluGlnAlaValValLeuAspCysThrLeuGly	58
Db	121	GGGCACTGCAGTGATCTGGGCGGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTG	180
Qy	59	AlaThrAlaAlaGlyProProThrArgValThrTrpSerLysAspGlyAspThrValLeu	78
Db	181	GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	240
Qy	79	GluHisGluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGlu	98
Db	241	GAGCACGACACTTACACCTGCTGGCCAAATGGTTCCCTGTGGCTGCTGCCAGCCAT	300
Qy	99	GlnGluAspSerAspAspGluGluAlaLeuArgLysTrpLysValThrGluGlySerTyr	118
Db	301	CCCAATGGCAGTGACGAGT--CAGTCCCTGAGGCTGTGGG-GTCAATGAGGCACTAT	357
Qy	119	SerCysLeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValValLysLeu	138
Db	358	TCGTGCTAGCCACGCGCCCTCGAGTGTGCGCCAGCCAGAGCTGCTGCTGCTGCTG	417
Qy	139	AlaThrLeuGluAspPheSerLeuHisProGlnSerGlnLeuValGluGluAsnGlyThr	158
Db	418	GCCAGTCTCGCAGACTTCTCTGACCCCGAGTCTCAGACGGTGGAGGAGAACGGACA	477
Qy	159	AlaArgPheGluCysHisThrLysGlyLeuProAlaProIleLeuThrTrpGluLysAsp	178
Db	478	GCTCGCTTGTGTCGCCACATTTGAAGGGTGCAGCTCCCATCATTTACTTGGAGAGGAC	537
Qy	179	GlnValThrValProGluGluProArgLysLeuThrLeuProLysTrpLeuLeuGlnIle	198
Db	538	CAGGTGACATTCGCTGAGGAGCTCGGCTCATCGTCTTCCACCGCGCTCTTCAGATC	597
Qy	199	LeuAspValGlnAspSerAspAlaGlySerTyrArgCysValAlaThrAsnSerAlaArg	218
Db	598	CTGATGTTTACGAGAGTGTATGACAGGCGCCCTACCGCTGCGTGGCCACCAACTCAG	657
Qy	219	GlnArgPheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAlaThr	238
Db	658	CAGCACTTACGACGAGGAGCCCTACTCATGTGTGGCCACAGAGTTTCTTGGCGTCCAC	717
Qy	239	ArgGlyGlnAspValValLeuAlaAlaProGluAsnThrThrValValSerGlyGln	258
Db	718	AGGGCGCAGGAGCGTGGTCATTGTGGCAGCCCGCAGAGAACCCACAGTGTGTCTGGC	777
Qy	259	AsnValValMetGluCysValAlaSerAlaAspProThrProPheValSerTrpValArg	278
Db	778	AGTGTGTGTGTAATGTGTGGCTGACGTGACCCCGCTTTTGTGTCTGTGGGTCCGA	837

Qy	279	GlnAspGlyLysProIleSerThrAspValIleValLeuGlyArgThrAsnLeuLeuLeu	298
Db	838	---GACGGAGAGCCCATCTCCACAGATGTATCTCTGGCGCCACCAACCTACTTAAT	894
Qy	299	AlaSerAlaGlnProArgHisSerGlyValTyrValCysArgAlaAsnLysProLeuThr	318
Db	895	GCCAAOCGCGAGCCCTGGCAGCTCCGCGCTCTATGTCTGCCGCGCCAAACAGCCCGC	954
Qy	319	ArgAspPheAlaThrAlaAlaGluLeuArgVal---LeuAlaAlaProAlaIleSer	337
Db	955	CGGACTTCCCACTGCGAGCCGCTGAGCTCCGTGTGTCTAGCGGCTCCCGCATCACT	1014
Qy	338	GlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArgAla	357
Db	1015	CAGCGCCCGGAGCGCTGCGCGAGCGCGCGGCGAGCAGCGCGCTCTGTGTGCGCGCG	1074
Qy	358	SerGlyGluProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgProAsn	377
Db	1075	TCGGGGGAGCGCGCGCGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1134
Qy	378	GlyArgValLysValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeuGln	397
Db	1135	GGCGCGCTCAGGTCCAGGGCGCGGTGGCGCTGTGTATCATCACAGATCGGCTGCG	1194
Qy	398	AspAlaGlyTyrTyrGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaAla	417
Db	1195	GACCGCGCTACTACAGTGGCTGGCTGAGAACAGCGCGGGAATGGCGTGGCTGCG	1254
Qy	418	ProLeuAlaValValArgGluGlyLeuProSerAlaProThrArgValThrAlaThr	437
Db	1255	TCGCTGCGCGTGGTGGTGGCGAGGGGCTGCCAGAGCGCCACCGGGTCACTGCTAC	1314
Qy	438	ProLeuSerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGluGln	457
Db	1315	CCACTGAGCAGCTCCCGTGTGTGGTGGCTGGAGCGGCGCCAGATGACAGCGAG	1374
Qy	458	IleIleGlyPheSerLeuHisTyrGlnLysAlaArgGlyValAspAsnValGluTyrGln	477
Db	1375	ATCATCGCTTCTCTCTCCACTACAGAGGACAGCGGGTATGACAAATGTGGATACC	1434
Qy	478	PheAlaValAsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrAsp	497
Db	1435	TTTGCACTGTAACCAACGACACACAGAACTACAGGTTCGGGACCTGGAAACCAAC	1494
Qy	498	TyrGluPheTyrValValAlaTyrSerGlnLeuGlyValAspSerArgThrSerSerPro	517
Db	1495	TATGAGTTCACGTGTGGTGGCTACTCCAGCTGGAGCGCCAGCCCTCCACCCCGCA	1554
Qy	518	LeuValHisThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSerPro	537
Db	1555	CTGTGTCACACACTGGATGATGTCCCAGTGCAGCAGCCCGCTCTCCCTGTCCAGCCC	1614
Qy	538	AsnProSerAspIleArgValAlaTrpLeuProLeuProSerSerLeuSerAsnGlyGln	557
Db	1615	AACCTTCGACATCAGGGTGGGTGGTGGCTGCGCCCGCCAGCTGAGCAATGGGCG	1674
Qy	558	ValLeuLysTyrIleGlyLeuGlyValGluAspGlnValPheSerThrGlu	577
Db	1675	GTGGTGAAGTACAGATAGAAATACGGTTTGGGAAAGGAAGATCAGATTTTCTTACT	1734
Qy	578	ValProGlyAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyrArg	597
Db	1735	GTGCGAGGAATGAGACACAGCTTATGCTGAACTCTGCTTTCAGCCCAACAGGTGTAT	1794
Qy	598	ValArgIleSerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGlnHis	617
Db	1795	GTACGATTTCCGCTGTGTACAGCAGCGGCTTCGGGCGCCCTCCAGTGTGATGCAT	1854
Qy	618	ArgThrProGlyValHisGlnSerHisValProPheAlaProAlaGluLeuLysVal	637
Db	1855	AGACGCCCGAGTATGACACCAACGAGCCATGTCCCTTTTGGCTGCGAGAGTGAAG	1914
Qy	638	ArgAlaLysMetGluSerLeuValValSerTrpGlnProProHisProThrGlnIle	657

1915 CAGGCAAGAGTGGAGTCCCTGGTGTCTATGGCAGCCACCCCTCACCCACCCAGATC 1974
658 SerGlyTyrLeuValTyrPheGlyValGlyThrGluGluGluAlaAspGlyAspArg 677
1975 TCTGGCTACAACTATATGGCGGAGGTGGGGCTGAGGAGGAGGCAATGGGATCGC 2034
678 ProProGlyArgGlyAspGlnAlaTyrPheValGlyProValArgLeuLeuLysLys 697
2035 CTGCCAGGGGGCCGTGGAGACCAAGCTTGGGATGGGGCTGTCTCCGGCTCAAGAGAAA 2094
698 ValLysGlnTyrGluLeuThrGlnLeuValProGlyArgProTyrGluValLysLeuVal 717
2095 GTGAAGCAGTATGAGCTGACCCAGCTAGTCCCTGGCCGCTGTACGAGGTGAAGCTCGT 2154
718 AlaPheAenLysHisGluAspGlyTyrAlaAlaValTyrLysGlyThrGluLysAla 737
2155 GCTTTCAACAACATGAGATGGCTATGACAGCTGTGGAGGCGACAGACGAGAGGGCG 2214
738 ProThrProAspLeuProIleGlnArgGlyProProLeuProProAlaHisValHisAla 757
2215 CCGGCACACAGATGCTATCCAGAGGGGACACCCCTGCTCCAGGCCACGTCATGGC 2274
758 GluSerAenSerThrSerIleTyrLeuArgTyrLysLysProAspPheThrVal 777
2275 GAATCAACAGCTCCACATCATCTGGCTTCGGTGGAGGAGGAGGAGGATTTTCCACACATC 2334
778 LysIleValAenTyrThrValArgPheGlyProTyrGlyLeuArgAenAlaSerLeuVal 797
2335 AAGATTGTCAACTACATCTGGCTTCAGCCCTGGGGGCTTCAGGAATGCTCTCTGGTC 2394
798 ThrTyrTyrThrSerSerGlyGluAspIleLeuIleGlyLeuLysProPheThrLys 817
2395 ACCTATTATC---AGTTCTGGAGAGACATCTCTATTTGGCGCTTGAAGCCATTCACCCAAA 2451
818 TyrGluPheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValVal 837
2452 TACGAGTTTGCAGTGCAGTCTCACGGCTGGACATGGATGGGCTTTCGGCTCTGTGGTG 2511
838 GluArgSerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerPro 857
2512 GAGGCTTCCACCTGCTGACCGTCTCCACACCCCTCCAGACCTGCGACTGAGCCCTC 2571
858 LeuThrProSerThrValArgLeuHisTyrCysProProThrGluProAenGlyGluIle 877
2572 CTGACACGCTCCACGGTTGGCTGCTCACTGGTGGCTCCACAGAGCCCAACGGGAGATC 2631
878 ValGluTyrIleLeuLeuTyrSerAenAenHisThrGlnProGluHisGlnTyrThrLeu 897
2632 GTGGAGTATCTGATCTCTGACAGCAGCAACACACAGCAGGCTGAGCACACAGTGGACCTTG 2691
898 LeuThrThr-----GluGlyAenIlePheSerAlaGluValHisGlyLeuGluSerAsp 915
2692 CTCACACGCGAGGTGAGGGAGAAATCTTCTAGTCTCAGGTCCATGGCTGGAGAGGCGAC 2751
916 ThrArgTyrPhePheLysMetGlyAlaArgThrGluValGlyProGlyProPheSerArg 935
2752 ACTCGTACTTCTCAAGATGGGGCGGCACAGAGTGGAGCTGGGCTTCTTCCCGC 2811
936 LeuGlnAspValIleThrLeuGlnGluThrPheSerAspSerLeuAspValHisAlaVal 955
2812 CTGCAGGATGTGATCAGCTCCAGGAGAGGTGTGACACTCGCTGGACATGCACTCAGTC 2871
956 ThrGlyIleIleValGlyValCysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 975
2872 ACGGGATCATCTGCGGTGTCTGCTGGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2931
976 GlyLeuArgGlnSerHisArgGluAlaLeuProGlyLeuSerSerSerGlyThrPro 995
2932 GGCCTGGCGGCGACCCCGCACAGGGATCTCCAGGCTGTCTCTCCAGCCGCCACCC 2991
996 GlyAenProAlaLeuTyrThrArgAlaArgLeuGlyProProSerValProAlaHis 1015

2992 GGGATCCCGCGCTGTACTCTCAGAGCTCGGCTTGGCCCCCCCCAGCCCCCAGCTGCCCAT 3051
1016 GluLeuGluSerLeuValHisProArgProGlnAspTyrSerProProSerAspVal 1035
3052 GAATTGGAGTCCCTTGTGCACCCCTCCAGAGCTGGTCCCGCCAGCCCTCAGAGCTG 3111
1036 GluAspLysAlaGluValHisSerLeuMetGlyLysSerValSerAspCysArgGlyHis 1055
3112 GAGGACAGGCTGAACTGACAGCTTATGGGTGGCGGTGTTTCTGAAGCCCGAGTGCAC 3171
1056 SerLysArgLysIleSerTyrAlaGlnAlaGlyLysProAenTyrAlaGlySerTyrAla 1075
3172 TCCAAAGAAAGTAAGT---GCTCAACAGCGGCTGAGCTGGGCTGTTCTCTGGCA 3228
1076 GlyCysGluLeuProGlnGlySerGlyProArgProAlaLeuThrArgAlaLeuLeuPro 1095
3229 GCGTGTGAGTGGCCAG---GCAGGCCCGCGCTCTGACCCCGGCTCTGCTGCC 3285
1096 ProAlaGlyThrGlyGlnThrLeuLeuLeuAlaLeuValTyrAspGlyLysSer 1115
3286 CTTGCTGGAACTGGGCGAGAGCTGTTGCTGCAAGTTCTCTGCTCTGAT-----CAGGGC 3339
1116 AenGlyArgLysLysProSerProAlaCysAenGlnValGluAlaGluValIleVal 1135
3340 AATGGGAGGAGGAAGTCAACCCAGCTGCGAGAACAGGTGAGGCTGAATGCTATGTC 3399
1136 HisSerAspPheGlyAlaSerLysGlyCysProAspLeuHisLeuGlnAspLeuGluPro 1155
3400 CACTCTGACTTGTAGTCACTTAACGGGAACTTGAACCTCCATCTCCAGACCTGGAGCT 3459
1156 GluGluProLeuThrAlaGluThrLeuProSerThrSerGlyAlaValAspLeuSerGln 1175
3460 GAGGACCCCTGCTCCAGAGGCTCTGATCTCATCTCGGGTGTGGAGTCCAGGGCAG 3519
1176 GlyAlaAspTyrLeuGlyArgGluLeuGlyCysGlnProThrThrSerGlyProGlu 1195
3520 GGGGCGACCTTGGCTGGACAGGGAGTTGGAGGCTGAGCTGGCAGCCCGGGCCAGAC 3579
1196 ArgLeuThrCysLeuProGluAlaAlaSerAlaSerCysSerCysSerAspLeuGlnPro 1215
3580 AGACTTACCTGCTTGGCAGGAGGAGGAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 3639
1216 SerThrAlaIleGluAlaProGlyLysSerCysGlnProLysAlaLeuCysProLeu 1235
3640 GGGCAGGTGTAGAGGAGACCCCTGGAGATAGTGGCAGCTCAATCCCTCTGCTCTCT 3699
1236 ThrValSerProSerLeuProArgAlaProValSerSerAla 1249
3700 GGAGCCAGCCCGGCTGGCCAGATCCCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3741

RESULT 5

US-09-754-997A-3
; Sequence 3, Application US/09754997A
; Patent No. US20020102551A1
; GENERAL INFORMATION:
; APPLICANT: Salbaum, Michael J.
; TITLE OF INVENTION: NOPE Polypeptides, Encoding Nucleic
; TITLE OF INVENTION: Acids and Methods of Use
; FILE REFERENCE: P-NI 4552
; CURRENT APPLICATION NUMBER: US/09/754,997A
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: US 60/174,496
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: US 60/205,789
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2796
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS

i LOCATION: (1)....(2796)

US-09-754-997A-3

Alignment Scores:

Pred. No.:	0	Length:	2796
Score:	4903.00	Matches:	932
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	74.24%	Indels:	0
DB:	10	Gaps:	0

US-09-754-997A-2 (1-1252) x US-09-754-997A-3 (1-2796)

Qy	22	GlyGluLeuProLeuGluThrValValLeuSerCysAspGluGlyProLeu	41
Db	1	GGGGAGCTGCCATTCGCCAGAGACAACTGTCAAGCTGAGCTGTGATGAGGGACCCCTG	60
Qy	42	GlnValIleLeuGlyProGluGlnAlaValValLeuAspCysThrLeuGlyAlaThrAla	61
Db	61	CAAGTGATCCTGGGCCCTGAGCAGGCTGTGTGTCTGGACTGCACCTTTGGGGGCTACAGCT	120
Qy	62	AlaGlyProProThrArgValThrTrpSerIysAspGlyAspThrValLeuGluHisGlu	81
Db	121	GCTGGGCTCCGACGAGGCTGACATGAGCAAGGATGAGACACTGTACTAGAGCATGAG	180
Qy	82	AsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGluGlnGluAsp	101
Db	181	AACCTGCACTTGTACCAATGGCTCCTGTGTGTCTGTCTCACCCTTAGGCAAGAGAC	240
Qy	102	SerAspAspGluGluAlaLeuArgIleTrpIysValThrGluGlySerTrpSerCysLeu	121
Db	241	AGCGATGATGAGGAGCTTTAGGATCTGGAAGTCACTGAGGCGAGCTATTCTCTGTCTG	300
Qy	122	AlaHisSerProLeuGlyValValAlaSerGlnValAlaValValLeuAlaThrLeu	141
Db	301	GCCCAAGCCCTAGGAGTGGTGGCCAGCCAGGTTGCTGTGTCTCAAGCTTGCCACATC	360
Qy	142	GluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThrAlaArgPhe	161
Db	361	GAAGACTTCTCTGCACCCGAGTCCAGATGTGTGGAGGAGACGGGACAGCAGCTTT	420
Qy	162	GluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAspGlnValThr	181
Db	421	GAATGCCACACCAAGGCTTCCAGCCCCCATCATTAATCTGGGNAAGGACGAGTGACC	480
Qy	182	ValProGluGluProArgLeuIleThrLeuProLysTrpLeuGluGlnIleLeuAspVal	201
Db	481	GTGCTTGAGGAGCCCGGCTCATCACTTCCCAAGTGCTCTCCAGATCCTAGATGTC	540
Qy	202	GlnAspSerAspAlaGlySerTrpArgCysValAlaThrAsnSerAlaArgGlnArgPhe	221
Db	541	CAGGACAGTGTGCAGGCTTCTACCGCTGCGTGGCCCAATTCAGCCCGCCACAGATTTC	600
Qy	222	SerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGln	241
Db	601	AGCCAGAGGCTTCGCTCACTGTGGCCCTCAGAGGCTCTTTGGAGCTACCAAGGGGCGAG	660
Qy	242	AspValValIleValAlaAlaProGluAsnThrThrValValSerGlyGlnAsnValVal	261
Db	661	GATGTGCTCATTTGTGGCAGCCCGAGAGAACACCGGTAGTGTCTGGACAGAAATGTAGTG	720
Qy	262	MetGluCysValAlaSerAlaAspProThrProPheValSerTrpValArgGlnAspGly	281
Db	721	ATGGAGTGGTGGCTCTGCTGACCCCAACCCCTTTGTGTCTTGGGTCCGACAGATGGA	780
Qy	282	LysProIleSerThrAspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSerAla	301
Db	781	AGCCATATCTCCAGGATGTCTGCTTCTGGCCGAGACCAATCTACTATCCCGAGGCG	840
Qy	302	GlnProArgHisSerGlyValTrpValCysArgAlaAsnLysProLeuThrArgAspPhe	321
Db	841	CAGCCTCGGCACTCTGGAGTCTATGTCTGCCGAGCCCAACAGCCCTCACGCGTGACTTC	900

Qy	322	AlaThrAlaAlaAlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGlu	341
Db	901	GCCACTGGGCTGCTGAGCTCCGAGTGTCTGCTGCCCGCCAGCATCTCGCAGGCACCCGAG	960
Qy	342	AlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluPro	361
Db	961	GCCTCTCGCGGACGCGGGCCAGACACGCGCGCTTGTGTGCGCGGCGTCCGCGGAGCCA	1020
Qy	362	ArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgValLys	381
Db	1021	CGGCCCCGCTGCACTGGCTGCAAGCGGANTCCGTTGCGACCCCAATGGGCGGCTCAG	1080
Qy	382	ValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTrp	401
Db	1081	GTGCAGGCGGCTGGCGCAGCTTGGTTCATCACTCAGATCGCTCGCAGGACGCTGGCTAC	1140
Qy	402	TyrGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaAlaProLeuAlaVal	421
Db	1141	TACAGTGGCTAGCAGAAAAACGCGCGGAACTGCCGTGTGCCCTGCGCCCTGGCGGTA	1200
Qy	422	ValValArgGluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSer	441
Db	1201	GTGTGCGCGAGGGGCTGCCAGCGCCCGCTCGGGTCAAGCCAGCCGCTGAGGAGC	1260
Qy	442	SerSerValLeuValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPhe	461
Db	1261	TCCTCTGTGTGTGGCTTGGAGCGGCTGAGTTGCACAGCGAGCAAAATCATTTGGCTTC	1320
Qy	462	SerLeuHisTrpGlnLysAlaArgGlyValAspAsnValGluTrpGlnPheAlaValAsn	481
Db	1321	TCTCTTCACTACCAAAAGCAAGGAGTGGACAAATGTGGAGTACCAAGTTTGCAGTAAC	1380
Qy	482	AsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThrArgTrpGluPheTrp	501
Db	1381	AATGACACCAAGAGCTGCGAGTTTCGGGACCTGGAACCCCAACCGGATTATAGTTCTAC	1440
Qy	502	ValValAlaTrpSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThr	521
Db	1441	GTGTGTGCTTCTCCAGCTGGGCGCCAGCCGAACCTCCAGCCCGAGCTGGTGCATACA	1500
Qy	522	LeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAsp	541
Db	1501	CTGAGCATGTCCCAGCGCAGCACCCCGAGCTTACCTTGTTCAGCCCGAACCCCTCGGAC	1560
Qy	542	IleArgValAlaTrpLeuProSerSerLeuSerSerGlnValLeuLysTrp	561
Db	1561	ATCAGGTTGGCATGGCTGCCCTGCCCTCCAGCTGAGCAATGACAGGCTGCTGAGTAC	1620
Qy	562	LysIleGluTrpGlyLeuGlyLysGluAspGlnValPheSerThrGluValProGlyAsn	581
Db	1621	AAGATAGATACGGTTTGGGGAAGGAAGATCAGGTTTTCTCCACCGAGGTGCTGGAAAT	1680
Qy	582	GluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTrpArgValArgIleSer	601
Db	1681	GAGACAACTTACGTTAAACTCACTTTCAGCAAAACAAAGTGTACCGAGTCCGGAATTC	1740
Qy	602	AlaGlyThrGlyAlaGlyTrpGlyValProSerGlnTrpMetGlnHisArgThrProGly	621
Db	1741	GCTGGCATGGCGCTGGCTATGGAGTCCCTTCTCAGTGGATGACAGCAGGACACCTGGT	1800
Qy	622	ValHisAsnGlnSerHisValProPheAlaProAlaGluLeuLysValArgAlaLysMet	641
Db	1801	GTGCACCAACAGAGCCATGTTCCCTTTGCCCCCTCGAGAAATGGAAGTGGGCGCAAGATG	1860
Qy	642	GluSerLeuValValSerTrpGlnProProHisProThrGlnIleSerGlyTrpLys	661
Db	1861	GAGTCCCTGGTGTGTCATGGCAGCGCCCTCCACCCAGCCAGATCTCTGGATACAA	1920
Qy	662	LeuTrpTrpGlyGluValGlyThrGluGluAlaAspGlyAspAspProGlyGly	681
Db	1921	CTCTACTGGGAGAGGTGGGAACAGAGAGGAGGAGATGGTGACCCCGCCCGAGGGGT	1980
Qy	682	ArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLysLysValLysValLysGlnTrp	701

Db 1981 CGTGGAGATCAAGCTTGGAGCGTGGGGCCGTCGCTGAAGAAGAAAGTGAAGCAGTAT 2040
Qy GluLeuThrGlnLeuValProGlyArgProTyrgluVallysLeuValAlaPheAenLys 721
Db 2041 GAATGACCCAGTATGCTCTGGCAGCGGTACAGGTGAGCTCTAGCTTTCACAA 2100
Qy HisGluAspGlyTyrAlaAlaValTyrLysGlyThrGluLysAlaProThrProAsp 741
Db 2101 CACGAGACGGCTACGCTGCTGTGTGAAGGGCAAGACGAGAGAGCGGCCACCGCAGAC 2160
Qy LeuProIleGlnArgGlyProProLeuProProAlaHisValHisAlaGluSerAenSer 761
Db 2161 CTGCTATTCAGAGGGGGCCACCGCTGCTCTGCCCCATGTCCACGAGAGTCAACAGC 2220
Qy SerThrSerIleTyrLeuArgTyrLysValProAspPheThrThrValLysIleValAsn 781
Db 2221 TCCACTTCCATTGGCTTGGTGGAGAGACGAGCTTACCACTCTGAGATTGTCAC 2280
Qy TyrThrValArgPheGlyProTyrGlyLeuArgAsnAlaSerLeuValThrTyrThr 801
Db 2281 TACACTGTACGCTTGGCCCTGGGGCTCAGGAATGCTTCCCTGGTCACTACTATACC 2340
Qy SerSerGlyGluAspIleLeuIleGlyGlyLeuLysProPheThrLysTyrGluPheAla 821
Db 2341 AGCTCTGGAGAGACATTTCTATTGGCGGCTGAAACCAATTTACCAAGTACGAGTTTGG 2400
Qy ValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSerThr 841
Db 2401 GTACGTCCACGGAGTGGATGGATGGGGCCCTTTGGCTCTGCTGAGAACCTCCACC 2460
Qy LeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrProSer 861
Db 2461 CTGCTGACCGGCTTCAACACCTCTCTGACCTGGCCCTGAGCCCTGACACCATCC 2520
Qy ThrValArgLeuHisTyrCysProProThrGluProAsnGlyGluIleValGluTyrLeu 881
Db 2521 ACCGTTTGGTTACATGGTGTCTCCCGCCAGCGGCCCAATGGTGGAGATTGTGGAGTATCTA 2580
Qy IleLeuTyrSerAenHisThrGlnProGluHisGlnTyrThrLeuLeuThrThrGlu 901
Db 2581 ATTCTCTACAGCAACCAACCAACCCAGCCGGAACACAGTGGGACCTGCTCCACAGAG 2640
Qy GlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePheLys 921
Db 2641 GGAAACATCTTCAGTCAGAGGTCATGGCTAGAGAGTACACTCGGTATTTCTTCAAG 2700
Qy MetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIleThr 941
Db 2701 ATGGAGCGCGCACAGAGTGGGGCCCTTTTCCGCTTGCAGGATGTGATTACT 2760
Qy 942 LeuGlnGluThrPheSerAspSerLeuAspValHis 953
Db 2761 CTGCAAGAGACATTCACAGACTCCTTGGATGTGCAC 2796

RESULT 6

US-10-239-663-28
; Sequence 28, Application US/10239663
; Publication No. US20030139572A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia, K.
; APPLICANT: Smith, Randall, F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kabnick, Karen
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50018
; CURRENT APPLICATION NUMBER: US/10/239,663
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09226
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/192,158

; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,668
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/200,166
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28
; LENGTH: 2223
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-239-663-28
Alignment Scores:
Pred. No.: 2,348-246 Length: 2223
Score: 3167.50 Matches: 620
Percent Similarity: 90.33% Conservative: 34
Best Local Similarity: 85.64% Mismatches: 54
Query Match: 47.96% Indels: 18
DB: 12 Gaps: 2
US-09-754-997A-2 (1-1252) x US-10-239-663-28 (1-2223)

Qy 17 LeuLeuSerAlaArgGlyGluLeuProLeuProGlnGluThrThrValLysLeuSerCys 36
Db 61 CTCCTCTCTATGCGAGGGAGCTGCTGTGCCCCCAGGAGACGACTGTGGAGCTGAGCTGT 120
Qy 37 AspGluGlyProLeuGlnValIleLeuGlyProGluGlnAlaValValLeuAspCysThr 56
Db 121 GGAGTGGGGCCACTGCAGTGATCTTGGGCCAGAGCGCTGCGAGTGTAACTGTATGAC 180
Qy 57 LeuGlyAlaThrAlaAlaGlyProThrArgValThrTrpSerLysAspGlyAspThr 76
Db 181 CTGGGGGCTGCTGCGCTGGACCCCGCCACAGGGTGACCTGGAGCAAGATGGGACACC 240
Qy 77 ValLeuGluHisGluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerPro 96
Db 241 CTGCTGGAGCACACCATTTACCTGTGCTGCCCAATGGTTCCCTGTGGCTGTCCAGCCA 300
Qy 97 LeuGluGlnGluAspSerAspAspGluGluAlaLeuArgIleTrpLysValThrGluGly 116
Db 301 CTAGCACCCCATGGCAGTGACAGGT--CAGTCCCTGAGGCTGTGGG--GTCAATTGAGGC 357
Qy 117 SerTyrSerCysLeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaVal 136
Db 358 AACTATTGCTGCTAGCCCGCCCGCTGGAGTGTGCTGCCAGCCAGACTGCTGTGCTC 417
Qy 137 LysLeuAlaThrLeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsn 156
Db 418 AAGCTTGGCCACACTGCGACAGTTCTCTCTGCACCCCGAGTCTCAGACGGTGGAGGAGAAC 477
Qy 157 GlyThrAlaArgPheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGlu 176
Db 478 GGGACAGCTCGCTTTGAGTGCCACATTGAAGGCTGCCAGCTCCCATCATTTACTTGGAG 537
Qy 177 LysAspGlnValThrValProGluGluProArgLeuIleThrLeuProLysTrpLeuLeu 196
Db 538 AAGGACCACTGACATGCTGAGGAGCTCTGGCTCATCTGTGCTTCCCAAGCGCTCTT 597
Qy 197 GlnIleLeuAspValGlnAspSerAspAlaGlySerTyrArgCysValAlaThrAenSer 216
Db 598 CAGATCTCGATGTTTTCAGGAGAGTATGAGGCGCCCTACCGCTGCGTGGCCACCACTCA 657
Qy 217 AlaArgGlnArgPheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGlu 236
Db 658 GCTCGCCAGCACTTCAGCCAGAGGCGCTTACTCAGTGTGGGCCACAGAGGCTCCCTGGCG 717
Qy 237 AlaThrArgGlyGlnAspValValIleValAlaProGluAenThrThrValValSer 256
Db 718 TCCACCGAGGGGCGAGACGCTGTCTATTGTGGAGCGCCCGAGAGACACCACTGCTGTCT 777
Qy 257 GlyGlnAsnValValMetGluCysValAlaSerAlaAspProThrProPheValSerTrp 276

```
Db 778 GGCACAGTGTGTGATGATGTGGCTCAGCTGACCCACCCCTTTTGTGTCTCG 837
Qy 277 ValArgGlnAspGlyLysProIleSerThrAspValIleValLeuGlyArgThrAsnLeu 296
Db 838 GTCCGACAAAGCGGAAGCCCATCTCCACAGATGTCTATGCTCTGGGCCGACCAACCTA 897
Qy 297 LeuIleAlaSerAlaGlnProArgHisSerGlyValTyrValCysArgAlaAsnLysPro 316
Db 898 CTAATTGCCAAGCGGAGCCCTGGCACTCCGGCGTCTATGTCTGCCGGCGCAACAGCCC 957
Qy 317 LeuThrArgAspPheAlaThrAlaAlaGlnLeuArgValLeuAlaAlaProAlaIle 336
Db 958 CGCACGGCGCACTTCGCGCACTGCAGCGCTGAGCTCGGTGTGTGGCGGTCTCCGCCATC 1017
Qy 337 SerGlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArg 356
Db 1018 ACTCAGGGCGCCGAGGGCGCTGTCCGGACGCGGGCGGACAGCGCGCTTCGTGTGCCG 1077
Qy 357 AlaSerGlyGluProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgPro 376
Db 1078 GCGTCGGGGGAGCCGCGCGCAGCGCTGGCTGGCTGCACAACGGGGCGCGCTGGGCCC 1137
Qy 377 AsnGlyArgValLysValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeu 396
Db 1138 AACGGCGCGTCAAGGTCAGGGCGCGGTGGCAGCGCTGGTCAATCACACAGATCGGCTG 1197
Qy 397 GlnAspAlaGlyTyrTyrGlnCysValAlaGlnAsnSerAlaGlyThrAlaCysAlaAla 416
Db 1198 CAGACCGCGCTACTACCAAGTGCCTGGCTGGAGAACAGCGCGGAATGGCGTGGCGTCC 1257
Qy 417 AlaProLeuAlaValValArgGluGlyLeuProSerAlaProThrArgValThrAla 436
Db 1258 GCGTCGCTGGCGGTGGTGGCGAGGGGTGCGCAGCGCGCCCAACCGGGTCACTGCT 1317
Qy 437 ThrProLeuSerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGlu 456
Db 1318 ACGCCACTAGCAGCTCGCTGTGTGTGGCTGGGAGCGCGCGGAGTGCACAGCGAG 1377
Qy 457 GlnIleIleGlyPheSerLeuHisTyrGlnLysAlaArgGlyValAsnValGluTyr 476
Db 1378 CAGATCATCGCTTCTCTCTCCACTACCAAGGACGCGGGCATGGACAATGTGGATAC 1437
Qy 477 GlnPheAlaValAsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThr 496
Db 1438 CAGTTTGAGTGAACACGACACACAGACTACAGTTTCGGGACCTTGGAAACCCACACA 1497
Qy 497 AspTyrGluPheTyrValValAlaTyrSerGlnLeuGlyAlaSerArgThrSerPro 516
Db 1498 GATTATGATGTTCTACGTGTGGCTTACTCCAGCTGGGAGCGCAGCCACCTCCACCCCA 1557
Qy 517 AlaLeuValHisThrLeuAspValProSerAlaAlaProGlnLeuThrLeuSerSer 536
Db 1558 GCACGTGGTGCACACTGGATGAT-----GGTAGGGCTCTGAACTCGCACTGGCGAGC 1611
Qy 537 ProAsnProSerAspIleArgValAlaTrpLeuProLeuProSerSerLeuSerAsnGly 556
Db 1612 -----TTGGCGCTGAGCAATGGG 1629
Qy 557 GlnValLeuLysTyrLysIleGluTyrGlyLeuGlyLysGluAspGlnValPheSerThr 576
Db 1630 CAGGTGGTGAAGTACAAGTACAGTACGTTTGGGAAGGAAGATCAGATTTTCTCTACT 1689
Qy 577 GluValProGlyAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyr 596
Db 1690 GAGGTGCGAGGAAATGAGACACAGCTTATGCTGAACCTCGCTTCAGCCAAACAGGTGTAT 1749
Qy 597 ArgValArgIleSerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGln 616
Db 1750 CGAGTACGAGTTCGGCTGGTACAGACCGCGCTTCGGGGGCCCTCCACGTGGATGAT 1809
Qy 617 HisArgThrProGlyValHisAsnGlnSerHisValProPheAlaProAlaGluLeuLys 636
Db 1810 CACAGGACCGCCAGTATGCACACAGAGGCCATGTCCCTTTTGGCCCTGCAGATGTGAG 1869
```

```
Qy 637 ValArgAlaLysMetGluSerLeuValValSerTrpGlnProProHisProThrGln 656
Db 1870 GTGCAGGCAAAAGTGGAGTCCCTGGTCTGTCTATGGCAGCCACCCCTCACCACCCAG 1929
Qy 657 IleSerGlyTyrLysLeuTyrTrpGlyGluValGlyThrGluGluGluAlaAspGlyAsp 676
Db 1930 ATCTCTGGCTTACAAACTATATTGGCGGAGGTGGGGCTGAGGAGGAGCCCAATGGCGAT 1989
Qy 677 ArgProProGlyGlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLysLys 696
Db 1990 CGCTCCGAGGGGCGCGTGGAGCAGGCTTGGGATGTGGGCTGTCCGGCTCAAGAAG 2049
Qy 697 LysValLysGlnTyrGluLeuThrGlnLeuValProGlyArgProTyrGluValLysLeu 716
Db 2050 AAGTGAAGCAGTATGAGCTGACCCAGCTAGTCCCTGGCGCGCTGTACGAGGTGAAGCTC 2109
Qy 717 ValAlaPheAsnLysHisGluAspGlyTyrAlaAlaValTrpLysGlyLysThrGluLys 736
Db 2110 GTGGCTTTCAACAAACATGAGGATGGCTATGCAGCAGTGTGGAAAGGCAAGCGGAGAAG 2169
Qy 737 AlaProThrPro 740
Db 2170 GCGCCGGCACCA 2181
```

RESULT 7

```
US-10-231-353-7
; Sequence 7, Application US/10231353
; Publication No. US20030023064A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20030023064A1el Human Phosphatases and Polynucleotides E
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/10/231,353
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: US/09/877,730
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 3453
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-231-353-7
```

```
Alignment Scores:
Pred. No.: 2,05e-131 Length: 3453
Score: 1747.50 Matches: 397
Percent Similarity: 54.95% Conservative: 169
Best Local Similarity: 38.54% Mismatches: 381
Query Match: 26.46% Indels: 83
DB: 14 Gaps: 18
```

US-09-754-997A-2 (1-1252) x US-10-231-353-7 (1-3453)

```
Qy 9 GlyLeuLeuValLeuThrPheCysLeuLeuSerAlaArgGlyGluLeuProLeuProGln 28
Db 43 GGGATGCTGCTCGCGCGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 96
Qy 29 GluThrThrVal---LysLeuSerCysAspGluGlnProLeuGlnValIleLeuGlyPro 47
Db 97 GTGTGGTGTCTTAGCGAAGTCTTTTGTAAAGAACCCACAGGATGTAACTGTCAACAAGA 156
Qy 48 GluGlnAlaValValLeuAspCysThrLeuGlyAlaThrAlaAlaGlyProProThrArg 67
Db 157 AAGGACCCAGTCTGTTTGTAGATTGCCAGGCTCACGAGAAAGTT-----CCTATTAAAG 207
```


Db 2257 ATTAACATACACCATCCGCTGTAATCTCTGTGGCTGCAGAAATGCTTCTTGGTTCTGTAC 2316
 Qy 800 TyrThrSerSerGlyGluAaplleLeulleGlyLeuLysProPheThrLysGlyGlu 819
 Db 2317 CTTCAAAATCAGAAATCTCATGTTGGTTCAAGGCTTAGAACCAACCAATACGAA 2376
 Qy 820 PheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArg 839
 Db 2377 TTTGCCGCTCGATACATGATGATCAGCTTCCAGTCTTGGAGCCCTGTAGTCTACCAT 2436
 Qy 840 SerThrLeuProAspArgProSerThrProSerAspLeuArgLeuSerProLeuThr 859
 Db 2437 TCTACTCTTCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2496
 Qy 860 ProSerThrValArgLeuHisThrCysProProThrGluProAsnGlyGluIleValGlu 879
 Db 2497 GATGACACTGCCCTGGTTCTTGGAAACCCCTGATGCCAGAAAAGTGTGGACCCG 2556
 Qy 880 TyrLeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTrrThrLeuThr 899
 Db 2557 TATATCTATATATGATCTAGGAGCCCTGGATTGAGGAGAGTGGCAGGCTTACAC 2616
 Qy 900 ThrGluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhe 919
 Db 2617 CGTGAAGGGGCAATAACCATGGCTTGTCTAGAAAACCTTGGTAGCAGGAAATGTGTACAT 2676
 Qy 920 PheLysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspVal 939
 Db 2677 GTCAAGATATCTGCATCAATGAGGTGGGAGAGGACCCCTTTTCAAAATCTGTGGAGCTG 2736
 Qy 940 IleThrLeuGlnGluThrPheSerAspSer 949
 Db 2737 GCAGTACTTCCAAAGGAACTCTGTAATCAATCAGAGGCCAAGCGTTAGATTCTGCT 2796
 Qy 950 959
 Db 2797 GATGCCAAAGTTTATTGAGGATATTACCATCTGCAGCAAAAATCAATGACTGGCAATGCT 2856
 Qy 960 ValGlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGln 979
 Db 2857 GTAGGTGTGGCAGTGGCTTGCATCTGATCTCTCATCTGTGTCTCTGTATATACCGA 2916
 Qy 980 SerSerHisArgGluAla- 987
 Db 2917 AGTAAGCCAGGAATCATCTGCTTCCAGAGCCGACAGATGGAATCAACAGTTACCT 2976
 Qy 988 GlyLeuSerSerGlyThrProGlyAsn 997
 Db 2977 CGTACCAGTGCCTCTTACGTAGTGGAAAT 3006

RESULT 8

US-10-231-353-1
 ; Sequence 1, Application US/10231353
 ; Publication No. US20030023064A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Walke, D. Wade
 ; APPLICANT: Scoville, John
 ; APPLICANT: Turner, C. Alexander Jr.
 ; APPLICANT: Friedrich, Glenn
 ; APPLICANT: Abuin, Alejandro
 ; APPLICANT: Zambrowicz, Brian
 ; APPLICANT: Sands, Arthur T.
 ; TITLE OF INVENTION: No. US20030023064A1el Human Phosphatases and Polynucleotides Encod
 ; FILE REFERENCE: LEX-0189-USA
 ; CURRENT APPLICATION NUMBER: US/10/231.353
 ; PRIOR FILING DATE: 2002-08-28
 ; PRIOR APPLICATION NUMBER: US/09/877,730
 ; PRIOR FILING DATE: 2001-06-08
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 3210
 ; TYPE: DNA

ORGANISM: homo sapiens
 US-10-231-353-1

Alignment Scores:
 Pred. No.: 2,256-131 Length: 3210
 Score: 1746.50 Matches: 397
 Percent Similarity: 54.95% Conservative: 169
 Best Local Similarity: 38.54% Mismatches: 381
 Query Match: 26.45% Indels: 83
 DB: 14 Gaps: 18

US-09-754-997A-2 (1-1252) x US-10-231-353-1 (1-3210)

Qy 9 GlyLeuLeuValLeuThrPheCysLeuLeuSerAlaArgGlyGluLeuProLeuProGln 28
 Db 43 GGGATGCTGCTCGCGCGCTCTCTGCTCTCTGCTGCTCAGT-----CCTTTGCCAGGA 96
 Qy 29 GluThrThrVal---LysLeuSerCysAspGlyProLeuGlnValLeuLeuGlyPro 47
 Db 97 GTGTGTGCTTTAGCGAATCTGCTTTTGTAAAGAACACACAGGATGTAACTGTACACAGA 156
 Qy 48 GluGlnAlaValValLeuAspCysThrLeuGlyAlaThrAlaAlaGlyProProThrArg 67
 Db 157 AAGGACCCAGTCTGTTTAGATTCCAGGCTCAGCGAGAGTT-----CCTATTAAAG 207
 Qy 68 ValThrTrpSerLysAspGlyAspThrValLeuGluHisGluAsnLeuHisLeuLeuPro 87
 Db 208 GTACATGTTGAAATATGGAGCAAAAATGTCTGMAAATAAACCGATCGAGGTTCTTCT 267
 Qy 88 AsnGlySerLeuTrpLeuSerSerProLeuGluGlnGluAspSerAspGluGluAla 107
 Db 268 AACGGCTCTTTATACATCAGT-----GAGGTGGAAGGAGGAGGAGGAGGAGTCC 318
 Qy 108 LeuArgIleTrpLysValThrGluGlySerTyrSerCysLeuAlaHisSerProLeuGly 127
 Db 319 -----GATGAAGATTTTATCAGTGTCTGCAATGAACAAATATGGA 360
 Qy 128 ValValAlaSerGlnValAlaValValLysLeuAlaThrLeuGluAspPheSerLeuHis 147
 Db 361 GCATTTTGTAGTCAAAAGCTCATCTGCTTATCACTATTCTTCATTTGAGTTCAGCAG 420
 Qy 148 ProGluSerGlnIleValGluGluAsnGlyThrAlaArgPheGluCysHisThrLysGly 167
 Db 421 CCAATTTCCATCAGGTCCAGAGGCGAGGTGGAGTGTCTGATTTCATGCAAGATTTTCATCC 480
 Qy 168 LeuProAlaProIleIleThrTrpGluLysAspGlnValThrValPro---GluGluPro 186
 Db 481 CACCCTCTCGACGTATACATGGGAGTTCAATCGGACAACTCTACCTATGACTATGGAC 540
 Qy 187 ArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuAspValGlnAspSerAspAla 206
 Db 541 AGATTAATCTCCCTACCAACAGGAGTATTGCAATCTATGATGTTCAGCCAAAGGATCT 600
 Qy 207 GlySerTyrArgCysValAlaThrAsnSerAlaArgGlnArgPheSerGlnGluAlaSer 226
 Db 601 GGAATATTCTGTTGCTGTCACCTGTAGCCACCGACGCTAAAGTATGAGAGGCTCG 660
 Qy 227 LeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAspValIleVal 246
 Db 661 CTAATCTGTGATT-----CCAGCTTAAGGAGTCAAAATCTCTCCACACACCAACCAATTTATA 714
 Qy 247 AlaAlaProGluAsnThrThrValValSerGlyGlnAsnValValMetGluCysValAla 266
 Db 715 GCAGGTCCAGAACATACACATCTCTTCATCAGACTGTAGTTTGGATATGATGATGCC 774
 Qy 267 SerAlaAspProThrProPheValSerThrValArgGlnAspGlyLysProIleSerThr 286
 Db 775 ACAGGAATCCCAAAACCAATCATTTCTTGGAGCGCTTGTATCATCACAATCCATTGATGTC 834
 Qy 287 ---AspValIleValLeuGlyArgThrAsnLeuLeuAlaSerAlaGlnProArgHis 305
 Db 835 TTTAATACTCGGCTACTTTGGAAATGGTAATCTCATGATATCTGATGTGAGCTACACAT 894


```
Db 2977 CGTACCAGTGCCTCTTAGCTAGTGTGAAAT 3006
RESULT 9
US-10-231-353-31
; Sequence 31, Application US/10231353
; Publication No. US20030023064A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20030023064A1 Human Phosphatases and Polynucleotides Encd
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/10/231,353
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: US/09/877,730
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 3874
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-231-353-31

Alignment Scores:
Pred. No.: 2,84e-131 Length: 3874
Score: 1746.50 Matches: 397
Percent Similarity: 54.95% Conservative: 169
Best Local Similarity: 38.54% Mismatches: 381
Query Match: 26.45% Indels: 83
DB: 14 Gaps: 18

US-09-754-997A-2 (1-1252) x US-10-231-353-31 (1-3874)
Qy 9 GlyLeuLeuValLeuThrPheCysLeuLeuSerAlaArgGlyGluLeuProLeuProGln 28
Db 233 GGGATGCTGTCGCCGGCTCTCTCTCTGTCGTGTCAGT-----CCTTTGCCAGGA 286
Qy 29 GluThrThrVal---LysLeuSerCysAspGluGlyProLeuGlnValLeuLeuGlyPro 47
Db 287 GTGTGGTCTTTAGCGAACTCTTTTGTAAAGAACCAACAGGATGTAACGTGTCAAGA 346
Qy 48 GluGlnAlaValValLeuApsCysThrLeuGlyAlaThrAlaAlaGlyProProThrArg 67
Db 347 AAGGACCCAGTCGTTTAGATTGCCAGGCTCACGAGAAAGTT-----CCTATTAAAG 397
Qy 68 ValThrTrpSerLysAspGlyAspThrValLeuGluHisGluAsnLeuLeuLeuPro 87
Db 398 GTCATGTTGTAATAATGGAGCAAAATGCTGAAATAAAGCGATCGAGGTCTTTCT 457
Qy 88 AsnGlySerLeuTrpLeuSerSerProLeuGlnGluAspSerAspAspGluGluAla 107
Db 458 AACGGCTTTTATACATCAGT-----GAGGTGGAAGCGACGCGGAGGAGCATCC 508
Qy 108 LeuArgIleTrpLysValThrGluGlySerTrpSerCysLeuAlaHisSerProLeuGly 127
Db 509 -----GATGAAGGATTTTATCAGTCTTGGCAATGAACAAATATGGA 550
Qy 128 ValValAlaSerGlnValAlaValValLeuAlaThrLeuGluApsPheSerLeuHis 147
Db 551 GCCATCTTAGTCAAAAGCTCATCTTGCCTTATCAACTATTTCTGCATTTGAAGTCCAG 610
Qy 148 ProGluSerGlnIleValGluGluAsnGlyThrAlaArgPheGluCysHisThrLysGly 167
Db 611 CCAATTTCCATGAGTCCAGAGGTGGAGTTCCTGATTTGATTCAGACAGATTTTCATCC 670
Qy 168 LeuProAlaProIleIleThrTrpGluLysAspGlnValThrValPro---GluGluPro 186
Db 671 CACCCTCTCGAGTCATAAACATGGGAGTTCAATCGGACAACTCTACCTATGACTATGGAC 730

187 ArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuApsValGlnApsSerAspAla 206
Db 731 AGGATAACTGCCCTACCAACAGAGTATTGCAGATCTATGATGTGACCCAAAGGGATTCT 790
Qy 207 GlySerTrpArgCysValAlaThrAenSerAlaArgGlnArgPheSerGlnGluAlaSer 226
Db 791 GGAATATTATCGTTGTATTGCTGCCACTGTAGCCACCGACGAGTAAAGATATGGAGGCTCG 850
Qy 227 LeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnApsValValIleVal 246
Db 851 CTAACGTGTGATT-----CCAGCTAAGGATCAAAATCTTCCACACACARCAATATA 904
Qy 247 AlaAlaProGluAsnThrThrValValSerGlyGlnAsnValValMetGluCysValAla 266
Db 905 GCAGGTCCACAGNACATACACATCTCTTCATCAGACTGTAGTTTGGNATOCATGGCC 964
Qy 267 SerAlaAspProThrProPheValSerTrpValArgGlnApsGlyLysProIleSerThr 286
Db 965 ACAGGAATCCCAACCAATCATTTCTTGAGCGCGCTTGATCACAATCCATTGATGTC 1024
Qy 287 ---AspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSerAlaGlnProArgHis 305
Db 1025 TTTAATACTCGGGTACTTGGAAATGTATCTCATGATCTCATGTCAGGCTACACAT 1084
Qy 306 SerGlyValTrpValCysArgAlaAsnLysProLeuThrArgAspPheAlaThrAlaAla 325
Db 1085 GCTGGAGTATATGTTGTGGGCCACTACCTCGGCACACGCACTTTACAGTTGCTATG 1144
Qy 326 AlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerArg 345
Db 1145 GCAACTTTAATCTGATTAGTCTCTCTCTTATTGTTGAATGGCCAGAAAGTTTAAACAAG 1204
Qy 346 ThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLeu 365
Db 1205 CCTCGAGCTGGCACTCTCGATTGTTGTGTCAGCGAAGGAATCCCTCTCCCAAGAG 1264
Qy 366 HisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgValLysValGlnGlyGly 385
Db 1265 TCATGTTGAAATAATGGAGGAATACATTCGATGTTAGATTAATTAATGTAC----- 1318
Qy 386 GlyGlySerLeuValIleThrGlnIleGlyLeuGlnApsAlaGlyTrpTrpGlnCysVal 405
Db 1319 AACAGTAAATGGTAAATTAACAGATATTTCCTGAAGATGATGCTATTATCATGTCATG 1378
Qy 406 AlaGluAsnSerAlaGlyThrAlaCysAlaAlaAlaProLeuAlaValValArgGlu 425
Db 1379 GCTGGAATAGCCAGGATCTATTATCTAGAGCCAGACTGACTGTAGTGTATGTCAGAA 1438
Qy 426 GlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerValLeu 445
Db 1439 GACAGACCCAGTCTCTCTATATATGATGTCGTAACCATGTCAGCTCAGCCATTCCT 1498
Qy 446 ValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTrp 465
Db 1499 TTAGCTCGGAGAGGCCACTTTTATAATTCAGACAAAGTCACTGCTATTCTGTACACTAC 1558
Qy 466 GlnLysAlaArgGlyValAspAsnValGlyTrpGlnPheAlaValAsnAspThrThr 485
Db 1559 ATGAAAGCAGAAAGTTTAAATAATGAAGAGTATCAAGTAGTCTCGGAAATGACACAACT 1618
Qy 486 GluLeuGlnValArgAspLeuGluProAsnThrAspTrpGluPheTrpValValAlaTrp 505
Db 1619 CATTATATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1678
Qy 506 SerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThrLeuAspAspVal 525
Db 1679 ATGCCAATGGAGCCAGCCAGAGTCTCTGACCATGTGACACAGATACTCTAGAGGATGT 1738
Qy 526 ProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAspIleArgValAla 545
Db 1739 CCCTCGAGACCTCTCGAAATTAGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1798
```


Db 100 CAGTGTGGCAATGAACAAATATGAGCCATTCTTAGTCAAAAGCTCATCTTGCCCTTA 159
Qy 139 AlaThrLeuGluAaspPheSerLeuHisProGluSerGlnIleValGluGluAenGlyThr 158
Db 160 TCAACTATTCTTGCATTGGAAGTCCAGCCCAATTTCCACTGAGGTCACGAGGTTGGAGTT 219
Qy 159 AlaArgPheGluCysHisThrLysGlyLeuProAlaProIleLeuThrTrpGluLysAasp 178
Db 220 GCTCGATTGTCATGCAAGATTTCATCCACCCTCTCCAGTCATAAACATGGAGTTCAAT 279
Qy 179 GlnValThrValPro---GluGluProArgLeuIleThrLeuProLysTrpLeuGln 197
Db 280 CGGCAAACTCTACCTATGACTATGAGCAGGATAACTGCCCTTACCACAGGAGTATTGACG 339
Qy 198 IleLeuAaspValGlnAaspSerAaspAlaGlySerTrpArgCysValAlaThrAenSerAla 217
Db 340 ATCTATGATGTGACCAAGAGATTCGGAATATTCGTTGATTTGCTGCCACTGTAGCC 399
Qy 218 ArgGlnArgPheSerGlnGluAaspSerLeuThrValAlaLeuArgGlySerLeuGluAla 237
Db 400 CACCGAGTAAAGTATGAGGCTCGTAACTGTGATT-----CCAGCTAAGGAGTCA 453
Qy 238 ThrArgGlyGlnAaspValIleValAlaAlaProGluAenThrThrValValSerGly 257
Db 454 AAATCCTTCCACACACCAACCAATATATAGCAGGTCCACAGAACATAACAACATCTCTTCAT 513
Qy 258 GlnAenValValMetGluCysValAlaSerLeuThrValAlaLeuArgGlySerLeuGluAla 277
Db 514 CAGACTGTAGTGTGGAAATGATGCCACAGGAATCCCAACCAATCTTCTTTGGAGC 573
Qy 278 ArgGlnAaspGlyLysProIleSerThr---AaspValIleValLeuGlyArgThrAenLeu 296
Db 574 GCGCTTGATCACAATCCATTGATGTCTTAAATACTCGGTACTTGGAAATGGTAATCTC 633
Qy 297 LeuIleAlaSerAlaGlnProArgHisSerGlyValTyrrValCysArgAlaAenLysPro 316
Db 634 ATGATATCTGATGTGAGCTACAACATGCTGGAGTATATGTTGTCGGGCCACTACCCCT 693
Qy 317 LeuThrArgAaspPheAlaThrAlaAlaGluLeuArgValLeuAlaAlaProAlaIle 336
Db 694 GGCACGCAACTTTACATGTTGCTATGGCACTTTAACTGATTAGTCTCTCTTCATTT 753
Qy 337 SerGlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArg 356
Db 754 GTTGAATGCCAGAAAGTTAAACAGGCTCGAGCTGCATCTCGATTGTTGTGTGAC 813
Qy 357 AlaSerGlyGluProArgProAlaLeuHisTrpLeuHisAaspGlyIleProLeuArgPro 376
Db 814 GCAGAAGGAATCCCTCTCCCAAGATGTGATGTTGAAATAATGGAAGGAAGATACATTGC 873
Qy 377 AasnGlyArgValIleValGlnGlyGlySerLeuValIleThrGlnIleGlyLeu 396
Db 874 AATGGTGAATTAATAATGTAC-----AACAGTAAATGGTAATTAACCAATATTTCCT 927
Qy 397 GlnAaspAlaGlyTyrrGlnCysValAlaGluAenSerAlaGlyThrAlaCysAlaAla 416
Db 928 GAAGATGATGCTATTATTCAGTGCATGCTGAGATAGCCCAAGGATCTATTATTATCTAGA 987
Qy 417 AlaProLeuAlaValValArgGluGlyLeuProSerAlaProThrArgValThrAla 436
Db 988 GCCAGACTGACTGTAGTGTGTCAGAGACAGACAGCCAGTGTCTCCATAATGTATCATGCT 1047
Qy 437 ThrProLeuSerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGlu 456
Db 1048 GAAACCACTGTCAGCTCAGCACTTTTTCAGCTGGGAGGCCCACTTTTATTAATTCAGAC 1107
Qy 457 GlnIleIleGlyPheSerLeuHisTyrrGlnLysAlaArgGlyValAaspAasnValGluTyrr 476
Db 1108 AAAGTCATTGCTATTCTGTACACTACATGAAGCAGAGAGTTTAAATATGAAGAGTAT 1167
Qy 477 GlnPheAlaAasnAasnAaspThrThrGluLeuGlnValArgAaspLeuGluProAasnThr 496
Db 1168 CAAGTAGTCATCGGAAATGACACAACTCATTTATATTGATGACTTAGAGCCTGCCAGC 1227

Qy 497 AaspTyrrGluPheTyrrValValAlaTyrrSerGlnLeuGlyAlaSerArgThrSerSerPro 516
Db 1228 AATTATCTTCTACATTTAGCATATATGCCAATCGGAGCCAGCCAGATGTCGACCAT 1287
Qy 517 AlaLeuValHisThrLeuAaspAaspValProSerAlaAlaProGlnLeuThrLeuSerSer 536
Db 1288 GTGACACAGAACTACTAGAGGATGTTCCTCGAGACCTCTCGAAATTAGTTTGAACAAGT 1347
Qy 537 ProAasnProSerAaspIleArgValAlaTrpLeuProLeuProSerSerLeuSerAasnGly 556
Db 1348 CGAAGTCCCACTCATATTTCTCTCTGCTGCCAATCCAGCCAAATATCGCGGGGC 1407
Qy 557 GlnValLeuLysTyrrLysIleGluTyrrGlyLeuGlyLysGluAaspGlnValPheSerThr 576
Db 1408 CAAAGTGTGCTGTATCGTGTGTTCTTCGCTTAAGTACTGAGAAATTCATTCCAAGTTCTG 1467
Qy 577 GluValProGlyAasnGluThrGlnLeuThrLeuAasnSerLeuGlnProAasnLysValTyrr 596
Db 1468 GAGCTCCCGGGACCAACGATGAGTACCTTTTGGAGGCTCGAAACCTGACAGTGTCTAC 1527
Qy 597 ArgValArgIleSerAlaGlyThrGlyAlaGlyTyrrGlyValProSerGlnTrpMetGln 616
Db 1528 CTGGTTTCGATTAATCTCTCTGCCACAGAGTGGGCTCGGAGAGTCATCAGTATGCACTTCA 1587
Qy 617 HisArgThrProGlyValHisAasnGlnSerHisValProPheAlaProAlaGluLeuLys 636
Db 1588 CATAGAGCGCCCAAAAGCT---ACAAGCGTGAAGCCCTTAAGTCTCCA---GAGTTGCAAT 1641
Qy 637 ValArg---AlaLysMetGluSerLeuValSerTrpGlnProProHisProThr 655
Db 1642 TTGGAGCCTCTGAACCTGTACCACTTCTGTGAGGTGGCAGCAAGATGTAGAGGACACA 1701
Qy 656 ---GlnIleSerGlyTyrrLysLeuTyrrGlyGluValGlyThrGluGluGluAlaAasp 674
Db 1702 GCTGCTATTGAGGCTACAGCTGTACTACAGGAAGAAGGCGCAGCAGAGAAAT----- 1755
Qy 675 GlyAaspArgProGlyGlyArgGlyAaspGlnAlaTrpAaspValGlyProValArgLeu 694
Db 1756 -----GGGCCCATTTTCTTG 1770
Qy 695 LysLysLysValLysGlnTyrrGluLeuThrGlnLeuValProGlyArgProTyrrGluVal 714
Db 1771 GATACCAAGGACCTACTCTATCTCTCAGTGGCTTAGACCCCAAGAAATAATATCATGTG 1830
Qy 715 LysLeuValAlaPheAenLysHisGluAaspGlyTyrrAlaAlaValTrpLysGlyLysThr 734
Db 1831 AGACTCTGCTGTACCAACACATAGACGATGGCTATCAGGCA-----GAT 1875
Qy 735 GluLysAlaProThrProAaspLeuProIleGlnArgGly-----ProProLeu 750
Db 1876 CAGACTGTACGACTCCAGGATCGGTCTGTGTCGTCATGTCGTCCTCTCCACCA 1935
Qy 751 ProProAlaHisValHisAlaGluSerAasnSerSerThrSerIleTrpLeuArgTrpLys 770
Db 1936 CCACCCCACTCTCTATGCAAGGCTAACACCTCATCTTCCATCTTCTCTGCACTGAGG 1995
Qy 771 LysProAaspPheThrValLysIleValAasnTyrrThrValArgPheGlyProTrpGly 790
Db 1996 AGCGCTGATTTCAACCGCTGCACAAATCATTAATACACCATCCGCTGTAAATCTGTTGGC 2055
Qy 791 LeuArgAasnAlaSerLeuValThrTyrrTyrrThrSerSerGlyGluAaspIleLeuIleGly 810
Db 2056 CTGAGATGCTTCTTTGGTTCTGTACTTCAAAACATCAGAAATCATCATGTTGGTTCAA 2115
Qy 811 GlyLeuLysProPheThrTyrrTyrrGluPheAlaValGlnSerHisGlyValAaspMetAasp 830
Db 2116 GGTCTAGAACCAACACCAAAATACGAATTTGCGCTTCGATTACATGATGATGATCAGCTTTC 2175
Qy 831 GlyProPheGlySerValValGluArgSerThrLeuProAaspArgProSerThrProPro 850
Db 2176 AGTCTTTGGAGCCCTGTAGTCTACCAATTTCTACTCTTCTTCAGAAAGCAGCAGGCCCA 2235

Score:	1472.00	Matches:	275
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	22.29%	Indels:	0
DB:	10	Gaps:	0

US-09-754-997A-2 (1-1252) x US-09-754-997A-5 (1-825)

Qy	978	ArgGlnSerSerHis	ArG	Gl	Ala	Leu	ProGly	Leu	SerSerGly	Thr	ProGly	Asn	997				
Db	1	CGACAAAGCTCCCA	CAGGAAAGCCCTT	CCCGGATTGTCTCT	CCTCAGGACACCCAGGAAC	60											
Qy	998	ProAlaLeu	TyrThr	Arg	Ala	Arg	Leu	GlyPro	SerVal	ProAla	AlaHis	Gl	Leu	1017			
Db	61	CCAGCGCTCTAC	ACNAGAGCTCGGCTT	GGGCTCCCAGTGTCCCTGCTGCCAT	AGATTG	120											
Qy	1018	Gl	SerLeu	ValHis	Pro	Arg	ProGln	AspTrp	SerPro	ProSer	AspVal	Gl	Asp	1037			
Db	121	GAGTCCCTCGT	GCATCTCTGTC	CCAGGATTGGTCCCAC	CACCCTCAGATGTGGAAGAC	180											
Qy	1038	Lys	Ala	Gl	ValHis	SerLeu	MetGly	CysVal	SerAsp	CysArg	GlyHis	Ser	Lys	1057			
Db	181	AAGGCTGAAGT	TACACACCTTAT	GGGTGGCAGTGTTCAGATTG	CCCGGGCCACTTCCAAAG	240											
Qy	1058	Arg	Lys	Thr	Ser	Trp	Ala	GlnAla	GlyGly	Pro	AsnTrp	Ala	GlySer	Trp	Ala	GlyCys	1077
Db	241	AGAAAGATCTCT	GGGCTCAGGC	AGGGGACCAACCTGGG	CAGGCTCTCTGGGCAGGCTGT	300											
Qy	1078	Gl	Leu	ProGln	GlySer	GlyPro	ArgPro	AlaLeu	ThrArg	AlaLeu	LeuPro	Pro	Ala	1097			
Db	301	GAGCTGCCCCAG	GGTAGTGGTCCAAG	CCCGCTCTGACCCGTGCTCTGCTGCCT	CCACGCG	360											
Qy	1098	Gly	Thr	GlyGln	ThrLeu	Leu	GlnAla	LeuVal	Tyr	AspGly	LysSer	Asn	Gly	1117			
Db	361	GGAACCCGGG	CAGACACTGCTCTG	CAAGCCCTGTGTATGACGCAT	TAAGAAGCAGCAGG	420											
Qy	1118	Arg	Lys	Leu	SerPro	AlaCys	Arg	AsnGln	ValGlu	AlaGlu	ValLys	ValHis	Ser	1137			
Db	421	AGAAAGAAGCG	CTCCCCAGCCCTG	CAGGAATCAGGTGGAAGCTG	AGGTCATTGTCCACTCC	480											
Qy	1138	Asp	Phe	GlyAla	SerLys	GlyCys	Pro	AspLeu	HisLeu	GlnAsp	LeuGlu	Pro	Glu	1157			
Db	481	GACTTCGGTGC	ATCCAAAGGATGTCCT	GACCTCCACCTCCAAGACCTTGG	AGCCAGAGGAA	540											
Qy	1158	Pro	Leu	ThrAla	Gl	ThrLeu	LeuPro	SerThr	SerGlyAla	ValAsp	LeuSer	GlnGly	Vala	1177			
Db	541	CCACTGACTGC	AGAGACTCTG	CCCTTCCAGCTCTGAGCTGTG	GAATCTGTCTCAAGAGACA	600											
Qy	1178	Asp	Trp	Leu	GlyArg	Gl	LeuGly	CysGln	ProThr	ThrSer	GlyPro	GluArg	Leu	1197			
Db	601	GACTGGCTGGG	CAGGAGCTGGGAGGGTGC	CAACCAACCACTAGTGGGCC	CAGAGAGGCTC	660											
Qy	1198	Thr	CysLeu	ProGlu	AlaAla	SerAla	SerCys	SerCys	SerAsp	LeuGln	ProSer	Thr	1217				
Db	661	ACCTGCTTGC	CAGAAGCAGCAGTGC	CTCTGCTCTCAGACCTTCC	AGCCAGCACT	720											
Qy	1218	Ala	Thr	Gl	GluAla	ProGly	LysSer	CysGln	ProLysAla	LeuCys	ProLeu	ThrVal	1237				
Db	721	GCTATAGAG	GAGGCCCTTGG	GAAGAAGCTGC	CAGCCCAAGCCCTGTGCTCTTAACATC	780											
Qy	1238	Ser	ProSer	LeuPro	ArgAla	ProVal	SerSer	AlaGln	ValPro	1252							
Db	781	AGCCCAAG	CCCTTCCAGGGCCCTGTCTCT	CTGCTGCTCAGGTCC	CC	825											

RESULT 13

US-10-231-353-5
; Sequence 5, Application US/10231353
; Publication No. US20030023064A1

; FACTICAL NO: US20
; GENERAL INFORMATION:
; GENERAL INFORMATION:

APPLICANT: Walke, D. Wade
APPLICANT: Scoville, John
APPLICANT: Turner, C. Ale
APPLICANT: Friedrich, Glee

Db	661	CTAACTGTGATT-----CGAGCTAAGGAGTCAAAATCTTCCACACACACCAACAATTATA	714
Qy	247	AlaAlaProGluAenThrThrValValSerGlyGlnAenValValMetGluCysValAla	266
Db	715	GCAGGTCACAGACATACACACATCTCTTCATCAGACTGTAGTTTTGGAAATGATGCC	774
Qy	267	SerAlaAspProThrProPheValSerTyrValArgGlnAenGlyLysProIleSerThr	286
Db	775	ACAGGAATCCAAACCATCAATTTCTTGGAGCGCTTGATCAAAATCCATTTGATGTC	834
Qy	287	---AspValIleValLeuGlyArgThrAenLeuLeuIleAlaSerAlaGlnProArgHis	305
Db	835	TTTAATACTCGGTACTTGGAAATGGTAATCTCATGATATCTCATGATCGAGGTACAAAT	894
Qy	306	SerGlyValTyrValCysArgAlaAenLysProLeuThrArgAspPheAlaThrAlaAla	325
Db	895	GCTGGAGTATATGTTTGGCGCCACTACCCCTGGCACACCAACTTTACAGTTGCTATG	954
Qy	326	AlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerArg	345
Db	955	GCACTTTAACTGATTAGTCTCTCTTCATTTGTTGATGGCCAGAAAGTTTAAACAGG	1014
Qy	346	ThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLeu	365
Db	1015	CCTCGAGCTGGCACTGCTCGATTGTTGTGTCTGAGCAGAAAGAAATCCCTCTCCCAAGATG	1074
Qy	366	HisTrpLeuHisAspGlyIleProLeuArgProAenGlyArgValLysValGlnGlyGly	385
Db	1075	TCATGGTTGAAATGGGAAGAGATACATTCGAATGGTAGAATTAANAATGATC-----	1128
Qy	386	GlyGlySerLeuValIleThrGlnIleGlyLeuGlnAenAlaGlyTyrTyrGlnCysVal	405
Db	1129	AACAGTAATTTGGTAATTAACAGATTAATTTCTGAGATGATGCTATTTATCAGTGCAATG	1188
Qy	406	AlaGluAenSerAlaGlyThrAlaCysAlaAlaProLeuAlaValValValArgGlu	425
Db	1189	GCTGAGATAGCCAAAGGATTAATTTATCTAGAGCCAGACTGACTGTAGTGTAGTCAGAA	1248
Qy	426	GlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerValLeu	445
Db	1249	GACAGACCCAGTGTCTCTATATGATGCTGAAACCATGTCTCAAGCTCAGCCATCTT	1308
Qy	446	ValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTyr	465
Db	1309	TTAGCCTGGGAGAGCCACTTTTATAATTCAGAAAGTCAATTCCTATCTGTACACTAC	1368
Qy	466	GlnLysAlaArgGlyValAspAenValGluTyrGlnPheAlaValAlaAenAspThrThr	485
Db	1369	ATGAAGCAGAGGTTTAAATAATGAAGATATCAAGTAGTCAATCGGAATGACACAAT	1428
Qy	486	GluLeuGlnValArgAspLeuGluProAenThrAspTyrGluPheTyrValValAlaTyr	505
Db	1429	CATTATATTATGATGACTTAGAGCTCCAGCAATTTATCTTTTACATTTGTACATAT	1488
Qy	506	SerGlnLeuGlyAlaSerArgThrSerProAlaLeuValHisThrLeuAspVal	525
Db	1489	ATGCCAATGGGAGCCAGATGCTGACCATGTGACACAGAAATACTCTAGAGATGAC	1548
Qy	526	ProSerAlaAlaProGlnLeuThrLeuSerSerProAenProSerAspIleArgValAla	545
Db	1549	CCACAG-----	1554
Qy	546	TrpLeuProLeuProSerSerLeuSerAenGlyGlnValLeuLysTyrLysIleGluTyr	565
Db	1554	-----	1554
Qy	566	GlyLeuGlyLysGluAspGlnValPheSerThrGluValProGlyAenGluThrGlnLeu	585
Db	1554	-----	1554
Qy	586	ThrLeuAenSerLeuGlnProAenLysValTyrArgValArgIleSerAlaGlyThrGly	605
Db	1554	-----	1554
Qy	606	AlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrProGlyValHisAenGln	625
Db	1554	-----	1554
Qy	626	SerHisValProPheAlaProAlaGluLeuLysValArgAlaLysMetGluSerLeuVal	645
Db	1554	-----	1554
Qy	646	ValSerTrpGlnProProHisProThrGlnIleSerGlyTyrLysLeuTyrTrpGly	665
Db	1554	-----	1554
Qy	666	GluValGlyThrGluGluGluAlaAspGlyAspArgProGlyGlyArgGlyAspGln	685
Db	1554	-----	1554
Qy	686	AlaTrpAspValGlyProValArgLeuLysLysValLysGlnTyrGluLeuThrGln	705
Db	1554	-----	1554
Qy	706	LeuValProGlyArgProTyrGluValLysLeuValAlaPheAenLysHisGluAspGly	725
Db	1555	-----AGAAATATCATGTGAGACTCTCGCTTACAAACACATAGACGATGCC	1602
Qy	726	TyrAlaAlaValTrpLysGlyLysThrGluLysAlaProThrProAspLeuProIleGln	745
Db	1503	TATCAGGCA-----GATCAGACTGTGAGACTCCAGGATGGTCTCTGTT	1647
Qy	746	ArgGly-----ProLeuProProAlaHisValHisAlaGluSerAenSer	761
Db	1648	CGTGATCGCATGGTCCCTCTCCACCACCCACCACCATCTCTATGCGAAGGCTTAACACC	1707
Qy	762	SerThrSerIleTrpLeuArgTrpLysLysProAspPheThrThrValLysIleValAen	781
Db	1708	TCATCTTCCATCTCTCCGACTGAGAGGCTGCAATTCACCGCTGCACCAATCATTAAC	1767
Qy	782	TyrThrValArgPheGlyProTrpGlyLeuArgAenAlaSerLeuValThrTyrTyrThr	801
Db	1768	TACACCATCGCTGTAATCTCTGTGGCTGCGAGATGCTCTTTGGTTCTGTACTTCAA	1827
Qy	802	SerSerGlyGluAspIleLeuIleGlyLysLysProPheThrLysTyrGluPheAla	821
Db	1828	ACATCAGAACTCACATGTTGGTTCAAGTCTAGAACCAACCAACCAATACGAATTTGCC	1887
Qy	822	ValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSerThr	841
Db	1888	GTTGATTTACATGTGGATCAGCTTTCAGTCTCTTGAGCCCTGTAGTCTACCATTTCTACT	1947
Qy	842	LeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrProSer	861
Db	1948	CTTCCAGAACACACGAGCCGCCACCATGTTGGAGTAAAGTGACATTAATAGAGATGAC	2007
Qy	862	ThrValArgLeuHisTyrCysProProThrGluProAenGlyGluIleValGluTyrLeu	881
Db	2008	ACTGCCCTGGTTCTTGGAAACCCCTGATGCGCCCAAGAACAGTTGTGACCCGCTACTACT	2067
Qy	882	IleLeuTyrSerAenAenHisThrGlnProGluHisGlnTrpThrLeuLeuThrThrGlu	901
Db	2068	ATCTTATATGATCTAGGAAGGCTGGAATTCAGAGAGAGTGGCAGGCTTACACCGTGAA	2127
Qy	902	GlyAenIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePheLys	921
Db	2128	GGGCAATAACCATGGCTTCTGTAAGAACTTGTAGCAGGAATGTGTATCTGTCGACG	2187
Qy	922	MetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIleThr	941
Db	2188	ATACTGATCAATCAGGTGGAGAGGACCCCTTTCAAAATCTGTGTGAGCTGGCAGTA	2247
Qy	942	LeuGlnGluThrPheSerAspSer-----	949
Db	2248	CTTCCAAAGGAAACCTCTGTAATCAAAATCAGAGGCCCAAGCGTTTAGATTCTGCTGATGCC	2307

Db 1369 ATGAAGCAGAGGTTTAAATAATGAAGAGTATCAAGTAGTCATCGGAATGACACAACT 1428
QY 486 GluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPheTyrValValalaTyr 505
Db 1429 CATTATATTATGATGACCTAGAGCTGCCAGCATTTACTTTCTACATTTAGCATAT 1488
QY 506 SerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThrLeuAspVal 525
Db 1489 ATGCCAATGGGACGACGATGCTGACCATGTGACACAGAAATACTCTAGAGGATGAC 1548
QY 526 ProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAspIleArgValAla 545
Db 1549 CCCAGA----- 1554
QY 546 TrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLysTyrLysIleGluTyr 565
Db 1554 ----- 1554
QY 566 GlyLeuGlyLysGluAspGlnValPheSerThrGluValProGlyAsnGluThrGlnLeu 585
Db 1554 ----- 1554
QY 586 ThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgIleSerAlaGlyThrGly 605
Db 1554 ----- 1554
QY 606 AlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrProGlyValHisAsnGln 625
Db 1554 ----- 1554
QY 626 SerHisValProPheAlaProAlaGluLeuLysValArgAlaLysMetGluSerLeuVal 645
Db 1554 ----- 1554
QY 646 ValSerTrpGlnProProHisProHisProHisProHisProHisProHisProHis 665
Db 1554 ----- 1554
QY 666 GluValGlyThrGluGluGluAlaAspGlyAspArgProGlyGlyArgGlyAspGln 685
Db 1554 ----- 1554
QY 686 AlaTrpAspValGlyProValArgLeuLysLysValLysValLysGlnTyrGluLeuThrGln 705
Db 1554 ----- 1554
QY 706 LeuValProGlyArgProTyrGluValLysLeuValAlaPheAsnLysHisGluAspGly 725
Db 1555 -----AGAAAAATATCATGTGAGACTCTGGCTTACAAACACATAGACGATGCC 1602
QY 726 TyrAlaAlaValTrpLysGlyLysThrGluLysAlaProThrProAspLeuProIleGln 745
Db 1603 TATCAGGCA-----GATCAGCTGTGACGACTCCAGGATGCGTGTCTGTT 1647
QY 746 ArgGly-----ProProLeuProAlaHisValHisValHisAlaGluSerAsnSer 761
Db 1648 CGTGATCGCATGGTCCCTCTCCACCACACCCACCATCTCTATGCGAAGGCTAACACC 1707
QY 762 SerThrSerIleTrpLeuArgTrpLysLysProAspPheThrThrValLysIleValAsn 781
Db 1708 TCATCTTCCATCTTCTGTCAGCTGAGGAGGCGCTGCATCTACCGCTGCACAAATCATTAAC 1767
QY 782 TyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTyrTyrThr 801
Db 1768 TACACCATCGCTGTAAATCTGTGGCTGCAGAAATGCTCTCTTGGTCTGTGATCTTCAA 1827
QY 802 SerSerGlyGluAspIleLeuIleGlyLysLeuLysProPheThrLysTyrGluPheAla 821
Db 1828 ACATCAGAAACTCATGTGGTGTGGTCTAGAGTCTAGAACCAACCAAAATACGAATTGCC 1887
QY 822 ValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSerThr 841
Db 1888 GTTCGATTACATGGGATCATGCTTTCCAGTCTTCCAGTCTTGGAGCCCTGTAGTCTACCATTTCTACT 1947

QY 842 LeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrProSer 861
Db 1948 CTTCCAGAAACACACAGCAGGCCACCACTTGGAGTAAAGTGACATTAATAGAGGATGAC 2007
QY 862 ThrValArgLeuHisTyrProProThrGluProAsnGlyGluIleValGluTyrLeu 881
Db 2008 ACTGCCCTGGTTCTTGGAAACCCCTGATGGCCAGAAACAGTGTGTGACCGCTATACT 2067
QY 882 IleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThrThrGlu 901
Db 2068 ATCTTATATGATCATCTAGGAAGCCCTGGATTGACAGGAGTGGCAGGTCTTACCGGTAA 2127
QY 902 GlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePheLys 921
Db 2128 GGGGCAATAACCATGGCTTTGCTAGAAAACTTGGTAGCAGGAAATGTGTACATTGTCAAG 2187
QY 922 MetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIleThr 941
Db 2188 ATATCTGCATCCAATGAGGTGGAGAGGCCCTTTTCAAAATTCGTGTGAGCTGCGAGTA 2247
QY 942 LeuGlnGluThrPheSerAspSer----- 949
Db 2248 CTTCCAAAGGAAACCTCTGAATCAATCAGAGGCCCAAGCGTTTAGATTCTGCTGATGCC 2307
QY 950 -----LeuAspValHisAlaValThrGlyIleIleValGly 961
Db 2308 AAAGTTTATTCAGGATATTACCATCTGCACCAAAAAATCAATGACTGGCATTCGTAGGT 2367
QY 962 ValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGlnSerSer 981
Db 2368 GTTGGCATAGCCTTGACCTGCATCCTCATCTGTGTCTCATCTGTGATATACCGAAGTAA 2427
QY 982 HisArgGluAla-----LeuProGlyLeu 989
Db 2428 GCCAGAAATCATCTGTTCCAGACGGCACAGAAATGGAACTCAACAGTTTACTCTGTACC 2487
QY 990 SerSerSerGlyThrProGlyAsn 997
Db 2488 AGTGCCTCTCTAGTAGTGAAAT 2511

RESULT 15

US-10-231-353-15
; Sequence 15, Application US/10231353
; Publication No. US20030023064A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: NO. US20030023064A1el Human Phosphatases and Polynucleotides Et.
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/10/231,353
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: US/09/877,730
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 2481
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-231-353-15

Alignment Scores:
Pred. No.: 8,66e-98 Length: 2481
Score: 1329.00 Matches: 312
Percent Similarity: 46.29% Conservative: 131
Best Local Similarity: 32.60% Mismatches: 276
Query Match: 20.12% Indels: 238

DB:	14	Gaps:	11
US-09-754-997A-2 (1-1252) x US-10-231-353-15 (1-2481)			
QY	79	GluHisGluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGlu	98
Db	7	GAATAAAGCGATCAGAGTTCTTCTAACGGCTCTTTATACATCAGT-----GAG	57
QY	99	GlnGluAspSerAspAspGluGluAlaLeuArgIleTrpLysValThrGluGlySerTyr	118
Db	58	GTGGAAGCGAGCGAGGAGCGAGTCC-----GATGAAGGATTTTAT	99
QY	119	SerCysLeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValValLysLeu	138
Db	100	CAGTGTCTGGCAATGAACAATATGAGCCATCTTAGTCAAAAAGCTCATCTCTGCTTA	159
QY	139	AlaThrLeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThr	158
Db	160	TCAACTATTCTGCAATTTGAAGTCCAGCCCAATTTCCACTGAGGTCCACGAAGTGGAGTT	219
QY	159	AlaArgPheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAsp	178
Db	220	GCTGATTTGCAATGCAAGATTTCATCCACCTCTCTGCAGTCATTAACATGGAGTTCAAT	279
QY	179	GlnValThrValPro---GluGluProArgLeuIleThrLeuProLysTrpLeuGln	197
Db	280	CGGACAACTCTACCTATGACTATGACAGAGATAAATGCCCTACCAACAGGAGTATGCAG	339
QY	198	IleLeuAspValGlnAspSerAspAlaGlySerTyrArgCysValAlaThrAsnSerAla	217
Db	340	ATCTATGATGTGAGCAAAAGGATTTCTGGAATATATCGTTGATTGTCGCCACTGTAGCC	399
QY	218	ArgGlnArgPheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAla	237
Db	400	CACCGACGTAAGAATGATGAGCGCTCGCTAACTGTGATT-----CCAGCTAAGGAGTCA	453
QY	238	ThrArgGlyGlnAspValValIleValAlaAlaProGluAsnThrThrValValSerGly	257
Db	454	AAATCCTTCCACACACCAACCAATATAGCAGGTCCACAGAACATAACAACATCTTCAT	513
QY	258	GlnAsnValValMetGluCysValAlaSerAlaAspProThrProPheValSerTrpVal	277
Db	514	CAGACTGTAGTTTGGAAATGATGCGCACAGGAATCCCAAAACCAATCATTTCTGGAGC	573
QY	278	ArgGlnAspGlyLysProLysThr---AspValIleValLeuGlyArgThrAsnLeu	296
Db	574	CGCCTTGATCACAAATCCATTCATGTCTTTAATATCTCGGGTACTTGGAAATGGTAACTCTC	633
QY	297	LeuIleAlaSerAlaGlnProArgHisSerGlyValTyrValCysArgAlaAsnLysPro	316
Db	634	ATGATATCTGTATGCTCAGGTACACATGCTGGAGTATATGTTTGTGGGCCCACTACCCCT	693
QY	317	LeuThrArgAspPheAlaThrAlaAlaGluLeuArgValLeuAlaAlaProAlaIle	336
Db	694	GGCACAGCAACTTTACAGTGTCTATGGCAACTTTAACTGTATTAGTCTCTCTTCATTT	753
QY	337	SerGlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArg	356
Db	754	GTGGAATGGCCAGAAAGTTTAAACAGGCCCTCGAGCTGGCACTGCTCGATTGTGTGTGAC	813
QY	357	AlaSerGlyGluProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgPro	376
Db	814	GCAGAGGAAATCCCTCTCCCAAGATGTCATGGTTGAAAATGGGAAGGAGATACATTCG	873
QY	377	AsnGlyArgValLysValGlnGlyGlyGlySerLeuValIleThrGlnIleGlyLeu	396
Db	874	AATGGTAGAATTTAAATGTAC-----AACAGTAAATTTGTAATTAACCAAGATTATTCCT	927
QY	397	GlnAspAlaGlyTyrTrpGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAla	416
Db	928	GAAGATGATGCTATTATATCATGTGCTGAGATAGCAAGGATCTATTTTATCTAGA	987
QY	417	AlaProLeuAlaValValArgGluGlyLeuProSerAlaProThrArgValThrAla	436
Db			
Db			
QY	988	GCCAGACTGACTGTAGTGAATGTCAAGACAGACCCAGTGTCTCCCTATAAATGTATCATGCT	1047
QY	437	ThrProLeuSerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGlu	456
Db	1048	GAACCAATGTCAAGCTCAGCCATCTTTTACCTTGGAGAGGCCACTTTATTAATTCAGAC	1107
QY	457	GlnIleIleGlyPheSerLeuHisTrpGlnLysAlaArgGlyValAspAsnValGluTyr	476
Db	1108	AAAGTCATTGCTATTCTGTACACTACATGAAGACAGAGGCTTTAAATAATGAAGAGTAT	1167
QY	477	GlnPheAlaValAsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThr	496
Db	1168	CAAGTAGTCATCGGAAATGACACAACTCATTTATTTATTTGATGATCTTAGAGCTGCCAGC	1227
QY	497	AspTyrGluPheTyrValValAlaTyrSerGlnLeuGlyAlaSerArgThrSerSerPro	516
Db	1228	AATTATACTTCTCATTTGTAGCATATATGCAATGGGAGCCAGCATGTCTGACCAT	1287
QY	517	AlaLeuValHisThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSer	536
Db	1288	GTGACACAGATACTCTAGAGGATGACCCAGA-----	1320
QY	537	ProAsnProSerAspIleArgValAlaTrpLeuProLeuProSerSerLeuSerAsnGly	556
Db	1320	-----	1320
QY	557	GlnValLeuLysTyrLysIleGluTyrGlyLeuGlyLysGluAspGlnValPheSerThr	576
Db	1320	-----	1320
QY	577	GluValProGlyAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyr	596
Db	1320	-----	1320
QY	597	ArgValArgIleSerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGln	616
Db	1320	-----	1320
QY	617	HisArgThrProGlyValHisAsnGlnSerHisValProPheAlaProAlaGluLeuLys	636
Db	1320	-----	1320
QY	637	ValArgAlaLysMetGluSerLeuValValSerTrpGlnProProHisProThrGln	656
Db	1320	-----	1320
QY	657	IleSerGlyTyrLysLeuTyrTrpGlyGluValGlyThrGluGluAlaAspGlyAsp	676
Db	1320	-----	1320
QY	677	ArgProProGlyGlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLysLys	696
Db	1320	-----	1320
QY	697	LysValLysGlnTyrGluLeuThrGlnLeuValProGlyArgProTyrGluValLysLeu	716
Db	1321	-----AGAAATATCATGTGAGACTC	1341
QY	717	ValAlaPheAsnLysHisGluAspGlyTyrAlaAlaValTrpLysGlyLysThrGluLys	736
Db	1342	CTGGCTTACACACACATAGACGATGGCTATCAGGCA-----GATCAGACT	1386
QY	737	AlaProThrProAspLeuProIleGlnArgGly-----ProProLeuProPro	752
Db	1387	GTCCAGCACTCCAGGATGCGTGTCTGTTCGTATGTCATGTCATGCTCCCTCCACACACCC	1446
QY	753	AlaHisValHisAlaGluSerAsnSerSerThrSerIleTrpLeuArgTrpLysLysPro	772
Db	1447	CACCATCTCTATGCGAAGGCTAACACCTCATCTTCCATCTTCTGCACTGGAGGAGGCT	1506
QY	773	AspPheThrThrValLysIleValAsnTyrThrValArgPheGlyProTyrGlyLeuArg	792
Db			

```
Db      1507 GCATTCACCGCTGCACAAATCAATTAACATCCGCTGTAATCCTGTTGGCCTGCAG 1566
Qy      793 AsnAlaSerLeuValThrTyrTyrThrSerSerGlyGluAspIleLeuIleGlyGlyLeu 812
Db      1567 AATGCTTCITTTGGTTCTGTACCTTCAACATCAGAACTCACAATGTGTGTTCAAGTCTA 1626
Qy      813 LysProPheThrIysTyrGluPheAlaValGlnSerHisGlyValAspMetAspGlyPro 832
Db      1627 GAACCAACACCAAAATACGAATTTGCCGTTTCGATTACATGTGATCAGCTTTCCAGTCCT 1686
Qy      833 PheGlySerValValGluArgSerThrLeuProAspArgProSerThrProProSerAsp 852
Db      1687 TGGAGCCCTGTAGTCTTACCACTTCTCTTCCAGAACACAGCAGCCCAACAGTTGGA 1746
Qy      853 LeuArgLeuSerProLeuThrProSerThrValArgLeuHisTrpCysProProThrGlu 872
Db      1747 GTAAAAGTGACATTATAGAGGATGACACTGCCCTGTTCTTGGAAACCCCTGATGGC 1806
Qy      873 ProAsnGlyGluIleValGluTyrLeuIleLeuTyrSerAsnAsnHisThrGlnProGlu 892
Db      1807 CCAGAAACAGTTGTGACCCGCTATCTATCTTATATGCATCTAGGAAGGCTGGATTGCA 1866
Qy      893 HisGlnTrpThrLeuLeuThrThrGluGlyAsnIlePheSerAlaGluValHisGlyLeu 912
Db      1867 GGAGAGTGGCAGGTCTTACACCGTGAAGGGGCAATAACCATGGCTTTGCTAGAAAACCTG 1926
Qy      913 GluSerAspThrArgTyrPhePheLeuMetGlyAlaArgThrGluValGlyProGlyPro 932
Db      1927 GTAGCAGGAATGTGTACATTGTCAAGATATCTGCATCCAAATGAGTGGAGAGGACCC 1986
Qy      933 PheSerArgLeuGlnAspValIleThrLeuGlnGluThrPheSerAspSer----- 949
Db      1987 TTTTCAAATTCGTGGAGCTGGCAGTACTTCCAAAGGAAACCTCTGAATCAAATCAGAGG 2046
Qy      950 -----LeuAspVal 952
Db      2047 CCCAAGCGTTTAGATTCTGCTGATGCCAAAGTTTATTTCAGGATATTACCATCTGGACCAA 2106
Qy      953 HisAlaValThrGlyIleIleValGlyValCysLeuGlyLeuLeuCysLeuAlaCys 972
Db      2107 AATCAATGACTGGCATGTCTGTAGGTGTGGCATGACCTTGACCTGCATCCTCATCTGT 2166
Qy      973 MetCysAlaGlyLeuArgGlnSerSerHisArgGluAla----- 985
Db      2167 GTTCTCATCTTGATATACCAAGTAAAGCCAGGAATCATCTGCTTCCAAGCGGCACAG 2226
Qy      986 -----LeuProGlyLeuSerSerSerGlyThrProGlyAsn 997
Db      2227 AATGGAACCTCAACAGTTACCTCGTCCAGTGCCTCTTGTAGTAGTGAAT 2277
```

Search completed: October 14, 2003, 09:45:00

Job time : 1026 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 14, 2003, 06:11:10 ; Search time 137 Seconds
(without alignments)
4033.662 Million cell updates/sec

Title: US-09-754-997A-2
Perfect score: 6504
Sequence: 1 MARADTGRLLVLTFCLLS.....CPLTVSPSLPRAPVSSAQVP 1252

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09754997/runat_14102003_061101_9736/app_query.fasta_1.1415
-DB=Issued_Patents_NA -QMT=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFTW=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09754997 SCGN 1.1.52 @runat_14102003_061101_9736 -NCPV=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	1747.5	26.5	3453	US-09-877-730-7
2	1746.5	26.4	3210	US-09-877-730-1
3	1746.5	26.4	3874	US-09-877-730-31
4	1691.5	25.6	2976	US-09-877-730-11
5	1691.5	25.6	3219	US-09-877-730-17
6	1384	21.0	2715	US-09-877-730-5
7	1384	21.0	2958	US-09-877-730-9
8	1329	20.1	2481	US-09-877-730-15
9	1329	20.1	2724	US-09-877-730-19
10	1144	17.3	2139	US-09-877-730-21
11	1144	17.3	2382	US-09-877-730-27
12	943	14.3	4608	US-09-041-886-24

13	943	14.3	4608	5	PCT-US94-05277-1	Sequence 1, Appli
14	909	13.8	1875	4	US-09-877-730-23	Sequence 23, Appli
15	781.5	11.8	1644	4	US-09-877-730-25	Sequence 25, Appli
16	781.5	11.8	1887	4	US-09-877-730-29	Sequence 29, Appli
17	647.5	9.8	1143	4	US-09-877-730-3	Sequence 3, Appli
18	638	9.7	6000	1	US-08-348-006B-6	Sequence 6, Appli
19	638	9.7	6000	2	US-08-800-825A-6	Sequence 6, Appli
20	638	9.7	6000	3	US-09-158-657-6	Sequence 6, Appli
21	638	9.7	6000	5	PCT-US94-10166-6	Sequence 6, Appli
22	634	9.6	4078	4	US-09-016-434-1132	Sequence 1132, Ap
23	592.5	9.0	909	4	US-09-877-730-13	Sequence 13, Appli
24	548.5	8.3	3888	4	US-08-506-296B-13	Sequence 13, Appli
25	540.5	8.2	3943	4	US-08-506-296B-27	Sequence 27, Appli
26	533	8.1	3774	2	US-08-341-843B-1	Sequence 1, Appli
27	533	8.1	3774	2	US-08-427-497B-2	Sequence 2, Appli
28	533	8.1	3774	2	US-08-427-497B-2	Sequence 2, Appli
29	522	7.9	5690	2	US-08-447-464-2	Sequence 2, Appli
30	522	7.9	5690	2	US-08-716-679-2	Sequence 2, Appli
31	514	7.8	3783	4	US-08-506-296B-20	Sequence 20, Appli
32	500.5	7.6	3189	2	US-08-427-497B-3	Sequence 3, Appli
33	500.5	7.6	5824	4	US-09-620-312D-72	Sequence 72, Appli
34	477	7.2	3991	4	US-08-506-296B-3	Sequence 3, Appli
35	443	6.7	6814	4	US-09-484-970B-66	Sequence 66, Appli
36	436.5	6.6	3360	1	US-08-408-093-5	Sequence 5, Appli
37	436.5	6.6	3360	1	US-08-408-420A-5	Sequence 5, Appli
38	436.5	6.6	3360	1	US-08-714-901-5	Sequence 5, Appli
39	436.5	6.6	3360	3	US-08-040-741-5	Sequence 4, Appli
40	387	5.9	2600	2	US-08-427-497B-4	Sequence 4, Appli
41	349.5	5.3	13857	4	US-09-620-312D-75	Sequence 75, Appli
42	324.5	4.9	3024	1	US-07-923-376-7	Sequence 7, Appli
43	320	4.8	4078	4	US-09-016-434-1120	Sequence 1120, Ap
44	320	4.8	4724	1	US-08-404-665-3	Sequence 3, Appli
45	320	4.8	4724	1	US-08-404-671-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-877-730-7
; Sequence 7, Application US/098777730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 3453
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-7

Alignment Scores:
Pred. No.: 4.42e-106 Length: 3453
Score: 1747.50 Matches: 397
Percent Similarity: 54.95% Conservative: 169
Best Local Similarity: 38.54% Mismatches: 381
Query Match: 26.46% Indels: 83
DB: 4 Gaps: 18

US-09-754-997A-2 (1-1252) x US-09-877-730-7 (1-3453)

QY 9 GlyLeuLeuValLeuThrPheCysLeuLeuSerAlaAargGlyGluLeuProLeuProGln 28
| | | | |
Db 43 GGGATGCTGCTCGCGGCTCTGCTCTGCTGCTCAGT-----CCTTGGCCAGA 96
| | | | |
QY 29 GluThrThrVal---LysLeuSerCysAaspGluGlyProLeuGlnValleLeuGlyPro 47
: : : : :
Db 97 GTGTGTGCTTTAGCGAAGCTGCTTTTGTAAAGAACACACAGGATGTAACGTGTCAACA 156
| | | | |
QY 48 GluGlnAlaValValLeuAaspCysThrLeuGlyAlaThrAlaAaGlyProProThrArg 67
: : : : :
Db 157 AAGGACCCAGTCTGTTTAGTTCAGGCTCAGGAGAAAGTT-----CCTATTAAG 207
| | | | |
QY 68 ValThrTrpSerLysAaspGlyAaspThrValLeuGluHisGluAasnLeuHisLeuLeuPro 87
: : : : :
Db 208 GTCACATGTTGAAATGGAGCAAAATGCTCGAAATAAACGGATCGAGGTCTTCTTCT 267
| | | | |
QY 88 AasnGlySerLeuTrpLeuSerProLeuGluGlnGluAaspSerAaspGluGluAla 107
: : : : :
Db 268 AACGGCTCTTTATACATCAGT-----GAGGTGGAAGGCAGGCGAGGAGCAGTC 318
| | | | |
QY 108 LeuArgIleTrpLysValThrGluGlySerTyrSerCysLeuAlaHisSerProLeuGly 127
| | | | |
Db 319 -----GATGAAGGATTTTATCAGTCTTGGCAATGAACAAATATGGA 360
| | | | |
QY 128 ValValAlaSerGlnValAlaValLysLeuAlaThrLeuGluAaspPheSerLeuHis 147
: : : : :
Db 361 GCCATCTTAGTCAAAAGCTCATCTGCTTATCACTATTTCTGCATTTGAGTCCAG 420
| | | | |
QY 148 ProGluSerGlnIleValGluGluAasnGlyThrAlaAargPheGluCysHisThrLysGly 167
| | | | |
Db 421 CCAATTTCCACTGAGGTCCACGAAGGTGGAGTTGCTCGATTTGCATGCAAGATTTTCATCC 480
| | | | |
QY 168 LeuProAlaProIleThrTrpGluLysAaspGlnValThrValPro---GluGluPro 186
| | | | |
Db 481 CACCCTCTCGATCATAAATGGAGTTCAATCGGACAACTCTACTATGACTATGGAC 540
| | | | |
QY 187 ArgLeuIleThrLeuProLysTrpLeuGlnIleLeuAaspValGlnAaspSerAaspAla 206
: : : : :
Db 541 AGGATAACTGCCCTACCAACAGAGATTTGCAGATCTATGATGTGACCCAAAGGATCT 600
| | | | |
QY 207 GlySerTyrArgCysValAlaThrAasnSerAlaAargGlnAargPheSerGlnGluAlaSer 226
| | | | |
Db 601 GGAATATTCGTTGTATTGCTGCCACTGTAGCCACCGACGTAAAGTATGGAGGCTCG 660
| | | | |
QY 227 LeuThrValAlaLeuAargGlySerLeuGluAlaThrAargGlyGlnAaspValValIleVal 246
| | | | |
Db 661 CTAACCTGTGATT-----CCAGCTAAGGAGTCAAAATCTCTCCACACACCAATATATA 714
| | | | |
QY 247 AlaAlaProGluAasnThrThrValValSerGlyGlnAasnValValMetGluCysValAla 266
| | | | |
Db 715 GCAGTCCACAGAACATACACATCTCTTCATCAGACTGTAGTTTGGAAATGATGGCC 774
| | | | |
QY 267 SerAlaAaspProThrProPheValSerTrpValAargGlnAaspGlyLysProIleSerThr 286
: : : : :
Db 775 ACAGGAATCCCAAAACCAATCATTTCTTGAGCGCGCTTCATCACAATCCATTGATGTC 834
| | | | |
QY 287 ---AaspValleValLeuGlyAargThrAasnLeuIleAlaSerAlaGlnProAargHis 305
: : : : :
Db 835 TTTAATACTCGGGTACTTGGAAATGTTAATCTCATGTATCTCATGTATCTGATGCTACACAT 894
| | | | |
QY 306 SerGlyValTrpValCysAargAlaAasnLysProLeuThrAargAaspPheAlaThrAlaAla 325
: : : : :
Db 895 GCTGGAGTATGTTTGTGGGCCCATCTACCCCTGGCACGCAACTTACAGTTGCTATG 954
| | | | |
QY 326 AlaGluLeuAargValLeuAlaProAlaIleSerGlnAlaProGluAlaLeuSerArg 345
: : : : :
Db 955 GCAACTTAACTGTATTAGTCTCTCTCTTATTGTTGAATGGCCAGAAAGTTTAAACAAG 1014
| | | | |
QY 346 ThrArgAlaSerThrAlaAargPheValCysAargAlaSerGlyGluProAargProAlaLeu 365
: : : : :
Db 1015 CCTCGAGCTGGCACTGCTCGATTTGTGTGTCAGCAGGAAGAAATCCCTCTCCCAAGATG 1074
| | | | |
QY 366 HistTrpLeuHisAaspGlyIleProLeuAargProAasnGlyAargValLysValGlnGly 385
| | | | |

Db 1075 TCATGGTTGAAAAATGGAAAGGAGATACATTCGAATGTTAGAAATGATAC----- 1128
| | | | |
QY 386 GlyGlySerLeuValIleThrGlnIleGlyLeuGlnAaspAlaGlyTyrTyrGlnCysVal 405
| | | | |
Db 1129 AACAGTAATTTGGTAATTAACCAAGATATTCTCTGAAGATGATGCTATTATTATCAGTGCATG 1188
| | | | |
QY 406 AlaGluAasnSerAlaGlyThrAlaCysAlaAlaProLeuAlaValValValAargGlu 425
: : : : :
Db 1189 GCTGAGATAGCCCAAGGATCTATTATTCTAGAGCCAGACTGACTGTAGTGTAGTCAGAA 1248
| | | | |
QY 426 GlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerValLeu 445
: : : : :
Db 1249 GACAGACCAGTGTCTCCCTATAATGTACATGCTGAAACCACTGCTCAAGCTCAGCCATTCTT 1308
| | | | |
QY 446 ValAlaIleTrpGluAargProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTyr 465
: : : : :
Db 1309 TTAGCTCTGGAGAGCCCACTTTATAATTCAGACAAAGTCATTGCTATTCTGTATCACTAC 1368
| | | | |
QY 466 GlnLysAlaAargGlyValAaspAasnValGluTyrGlnPheAlaValAasnAasnAaspThrThr 485
: : : : :
Db 1369 ATGAAGCAGAGGTTTAATAATGAAGATGATCAAGTAGTCAATCGGAATGACCAACT 1428
| | | | |
QY 486 GluLeuGlnValAargAaspLeuGluProAasnThrAaspTyrGluPheTyrValValAlaTyr 505
: : : : :
Db 1429 CATTTATATTATGATGACTTAGAGCTGCAGCAATTTATCTTCTCATTTGTAGCATAT 1488
| | | | |
QY 506 SerGlnLeuGlyAlaSerArgThrSerProAlaLeuValHisThrLeuAaspVal 525
: : : : :
Db 1489 ATGCCAATGGGAGCCAGCCAGATGCTCACCATTGTCAGACACAGAACTCTTAGAGATGTT 1548
| | | | |
QY 526 ProSerAlaAlaProGlnLeuThrLeuSerSerProAasnProSerAaspIleArgValAla 545
: : : : :
Db 1549 CCCCTGAGACCTCTCTGAAATTTAGTTTGACAGTCAAGTCCACCATGATATTCTCATCTCC 1608
| | | | |
QY 546 TrpLeuProLeuProSerSerLeuSerAasnGlyGlnValLeuLysTyrLysIleGluTyr 565
: : : : :
Db 1609 TGGCTGCCAATCCCAAGCAATATCGGGGGCCAAAGTGGTGTCTGTATCGTCTGCTTTC 1668
| | | | |
QY 566 GlyLeuGlyLysGluAaspGlnValPheSerThrGluValProGlyAasnGluThrGlnLeu 585
| | | | |
Db 1669 CCGCTTAAGTACTAGAAATTCATCAATCCAAAGTTCTGGAGCTCCCGGGGACCAACGATGATAC 1728
| | | | |
QY 586 ThrLeuAasnSerLeuGlnProAasnLysValTyrAargValAargIleSerAlaGlyThrGly 605
| | | | |
Db 1729 CTTTGTGAAGGCTCGAAACCTGACAGTGTCTACCTGTTTGGATTTACTCTGCCACCAGA 1788
| | | | |
QY 606 AlaGlyTyrGlyValProSerGlnTrpMetGlnHisAargThrProGlyValHisAasnGln 625
: : : : :
Db 1789 GTGGGGCTGGGAGAGTCATCAGTATGGACTTCACATAGGACGCCCAAGCT---ACAAGC 1845
| | | | |
QY 626 SerHisValProPheAlaProAlaGluLeuLysValAarg---AlaLysMetGluSerLeu 644
: : : : :
Db 1846 GTGAAGGCCCTTAAGTCTCCA---GAGTTGCAATTTGGAGCCCTCTGAACCTGTATCCACCAT 1902
| | | | |
QY 645 ValValSerTrpGlnProProProHisProThr---GlnIleSerGlyTyrLysValLeuTyr 663
: : : : :
Db 1903 TCTGTGAGTGGCAGCAAGATGTAGAGGACACAGCTGCTATTTCAGGGCTACAGCTGTATAC 1962
| | | | |
QY 664 TrpGlyGluValGlyThrGluGluAlaAaspGlyAaspArgProProGlyGlyArgGly 683
: : : : :
Db 1963 TACAAGGAAGAGGCGCAGGAGAAAT----- 1989
| | | | |
QY 684 AspGlnAlaTrpAaspValGlyProValAargLeuLysLysValLysValGlnTyrGluLeu 703
| | | | |
Db 1990 -----GGGCCCATTTTCTGGATACCAAGGACCTCTACTATATCTCTC 2031
| | | | |
QY 704 ThrGlnLeuValProGlyAargProTyrGluValLysLeuValAlaPheAasnLysHisGlu 723
: : : : :
Db 2032 AGTGGCTTAGACCCCAAGAAATATCATGTGAGACTCTCTGGCTTACCAACACATAGAC 2091
| | | | |
QY 724 AaspGlyTyrAlaAlaValTrpLysGlyLysThrGluLysAlaProThrProAaspLeuPro 743
| | | | |

```

Db 2092 GATGGCTATCAGGCA-----GATCAGACTGTCCAGCACTCCAGGATGCGTG 2136
Qy 744 lIeGlnAArgGly-----ProLeuPProAlaHisValHisAlaGluSer 759
Db 2137 TCTGTGCTGATCGCATGCTCTCTCCACCAACCCACCACTCTCTATCGGAGGCT 2196
Qy 760 AenSerSerThrSerIleThrLeuArgTrpLysLysProAspPheThrThrValLysIle 779
Db 2197 AACACCTCATCTTCCATCTCTCTGCACTGGAGAGGCTGCAATTCACCGCTGCACAAATC 2256
Qy 780 ValAenThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTrp 799
Db 2257 ATTAACTACACCATCGCTGTAATCTGTGCTGCTGCAAGATGCTCTTGTGTTCTGTAC 2316
Qy 800 TyrThrSerSerGlyGluAspIleLeuIleGlyGlyLeuLysProPheThrLysTrpGlu 819
Db 2317 CTTCAACATCAGAACTACATGTTGTTCAAGTCTAGAACCAACCAACCAATACGAA 2376
Qy 820 PheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArg 839
Db 2377 TTTGGCGTTCGATTACATCTGCATCAGCTTTCAGTCTTTCAGGCTTTCAGGCTTTCAGCAT 2436
Qy 840 SerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThr 859
Db 2437 TCTACTCTTCCAGAACGACACGACGCGCCACCACTGTTGGAGTAAAGTGAATATAGAG 2496
Qy 860 ProSerThrValArgLeuHisTrpCysProThrGluProAsnGlyGluValGlu 879
Db 2497 GATGACACTGCCCTGTTCTTGGAAACCCCTGATGGCCCAAGAACAGTGTGACCCGC 2556
Qy 880 TyrLeuIleLeuTrpSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThr 899
Db 2557 TATACTATCTTATGATCATCTAGGAGGCTGATGTTGACAGGAGTGGCAGGCTTTACAC 2616
Qy 900 ThrGluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTrpPhe 919
Db 2617 COTGAAGGGGCAATAACCACTGCTTGTAGAAAACCTTGTGACGAGGAATGTACATT 2676
Qy 920 PheLysMetGlyValArgThrGluValGlyProGlyProPheSerArgLeuGlnAspVal 939
Db 2677 GTCAGATATCTGCATCCATCAGGTGGAGGAGGACCCCTTTTCAATCTCTGGAGCTG 2736
Qy 940 IleThrLeuGlnGluThrPheSerAspSer----- 949
Db 2737 GCAGTACTTCCAAAGGAACCTCTGAATCAATCAATCAGAGCCCAAGGTTTAGATTCTGCT 2796
Qy 950 -----LeuAppValHisAlaValThrGlyIleIle 959
Db 2797 GATGCCAAAGTTTATTTCAGGATATTACCATCTGGACCAAAAATCAATGACTGGCATGCT 2856
Qy 960 ValGlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGln 979
Db 2857 GTAGGTGTGGCATAGCCCTGACCTGATCCTCATCTGTGTTCTCATCTTTGATATACCGA 2916
Qy 980 SerSerHisArgGluAla-----LeuPro 987
Db 2917 AGTAAAGCCAGGAATCATCTGCTTCCAGACGGCACAGATGGAATCAACAGTTACTT 2976
Qy 988 GlyLeuSerSerSerGlyThrProGlyAsn 997
Db 2977 CGTACCAGTGCCTCTTAGCTAGTGGAAAT 3006

```

RESULT 2

US-09-877-730-1

; Sequence 1, Application US/0987730

; Patent No. 6465632

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Scoville, John

; APPLICANT: Turner, C. Alexander Jr.

; APPLICANT: Friedrich, Glenn

; APPLICANT: Abuin, Alejandro

; APPLICANT: Zambrowicz, Brian

```

; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632e1 Human Phosphatases and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3210
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-1

```

Alignment Scores:

```

Pred. No.: 4,66e-106 Length: 3210
Score: 1746.50 Matches: 397
Percent Similarity: 54.95% Conservative: 169
Best Local Similarity: 38.54% Mismatches: 381
Query Match: 26.45% Indels: 83
DB: 4 Gaps: 18

```

US-09-754-997A-2 (1-1252) x US-09-877-730-1 (1-3210)

```

Qy 9 GlyLeuLeuValLeuThrPheCysLeuLeuSerAlaArgGlyGlyLeuProLeuProGln 28
Db 43 GGGATGCTGCTCCGCGGCTCTCTGCTCTGCTGCTCAGT-----CCTTTGCCAGGA 96
Qy 29 GluThrThrVal-----LysLeuSerCysAspGluGlyProLeuGlnValIleLeuGlyPro 47
Db 97 GTGTGCTGCTTTAGCAACTGCTCTTTGTAAGAAGAACCAACAGGATGTAACTGTCAAGA 156
Qy 48 GluGlnAlaValValLeuAspCysThrLeuGlyAlaThrAlaAlaGlyProProThrArg 67
Db 157 AAGGACCCAGTCGTTTAGATTGTCAGGCTCAGGAGAAAGTT-----CCTATTAAAG 207
Qy 68 ValThrTrpSerLysAspGlyAspThrValLeuGluHisGluAsnLeuHisLeuLeuPro 87
Db 208 GTCACATGTTGAAAAATGGAGCAAAATGCTGAAATAAAACGATCGAGGTCTTTCT 267
Qy 88 AsnGlySerLeuTrpLeuSerSerProLeuGlnGluAspSerAspAspGluAla 107
Db 268 AACGGCTCTTTATACATCAGT-----GAGGTGGAAGGCGAGGAGGAGCAGTCC 318
Qy 108 LeuArgIleThrLysValThrGluGlySerTrpSerCysLeuAlaHisSerProLeuGly 127
Db 319 -----GATGAAGGATTTTATCAGTCTCTGGCAATGAACAAATATGGA 360
Qy 128 ValValAlaSerGlnValAlaValValLysLeuAlaThrLeuGluAspPheSerLeuHis 147
Db 361 GCCATTCTTAGTCAAAAAGCTCATCTTGCCTTATCACTATTCTTGCATTGAGTCCAG 420
Qy 148 ProGluSerGlnIleValGluGluAsnGlyThrAlaArgPheGluCysHisThrLysGly 167
Db 421 CCAATTTCCACTCAGGTCCACGAGGTGGAGTGTCTGATTTGCATGCAAGATTTTCATCC 480
Qy 168 LeuProAlaProIleIleThrTrpGluLysAspGlnValThrValPro---GluGluPro 186
Db 481 CACCTCTCGCATCATCAACATGGGAGTTCAATCGACACACTCTACCTATGACTATGGAC 540
Qy 187 ArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuAspValGlnAspSerAspAla 206
Db 541 AGGATACTGCCCTTACCACAGGAGTATTGAGATCTATGATGTGACCAAGGAGATTCT 600
Qy 207 GlySerTrpArgCysValAlaThrAsnSerAlaArgGlnArgPheSerGlnGluAlaSer 226
Db 601 GGAATATTATGTTGTTGTTGCTGCCACTGTAGSCCCACCGAGCTAAAGATATGAGGCTCG 660
Qy 227 LeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAspValIleVal 246
Db 661 CTAACCTGTGATT-----CCAGCTAAGGAGTCAAAATCCTTCCACACACCAATTTATA 714

```



```

Db      2797 GATGCCAAAGTTTATTCAGATATTACCATCTGGACCAAAAATCAATGACTGGCATTTGCT 2856
Qy      960 ValGlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArgGln 979
Db      2857 GTAGGTGTTGGCATACGCTTGACCTGCATCTCTCATCTGTGTTCTCATCTTGATATACCGA 2916
Qy      980 SerSerHisArgGluAla-----LeuPro 987
Db      2917 AGTAAAGCCAGGAATCATCTGTTCCAAAGCGGCACAGATGGAACCTCAACAGTTTACCT 2976
Qy      988 GlyLeuSerSerGlyThrProGlyAsn 997
Db      2977 CGTACAGTGCCTCTTACTAGTGGAAAT 3006

RESULT 3
US-09-877-730-31
; Sequence 31, Application US/09877730
; Patent No. 645632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 645632el Human Phosphatases and Polynucleotides Encoding the
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 3874
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-31

Alignment Scores:
Pred. No.:      6,048-106      Length:      3874
Score:          1746.50      Matches:    397
Percent Similarity: 54.95%      Conservative: 169
Best Local Similarity: 38.54%      Mismatches: 381
Query Match:      26.45%      Indels:     83
DB:               4          Gaps:        18

US-09-754-997A-2 (1-1252) x US-09-877-730-31 (1-3874)
Qy      9 GlyLeuLeuValThrPheCysLeuLeuSerAlaArgGlyGluLeuProLeuProGln 28
Db      233 GGGATGCTGCTCGGCGCTCTCTGCTCTGCTGCTCAGT-----CCTTTGCCAGA 286
Qy      29 GluThrThrVal---LysLeuSerCysAspGluGlyProLeuGlnValIleLeuGlyPro 47
Db      287 GTGTGTGCTTTAGCGAAGCTGTTTGTGAAAAGAACACACAGATGTAACTGTCAAGA 346
Qy      48 GluGlnAlaValValLeuAspCysThrLeuGlyAlaThrAlaAlaGlyProThrArg 67
Db      347 AAGGACCCAGTCGTTTATGATTCGCCAGGCTCACGGAGAGT-----CCTATTAA 397
Qy      68 ValThrTrpSerLysAspGlyAspThrValLeuGluHisGluAsnLeuHisLeuLeuPro 87
Db      398 GTCATAGTTGGAATAATGGACAAAATGCTCGAAAATAAACCAGGATGAGGTTCTTTCT 457
Qy      88 AsnGlySerLeuTrpLeuSerSerProLeuGluGlnGluAspSerAspAspGluGluAla 107
Db      458 AACGGCTCTTTATACATCAGT-----GAGGTGGAAGCGCAGCGGAGGAGGAGTCC 508
Qy      108 LeuArgIleTrpLysValThrGluGlySerThrCysLeuAlaHisSerProLeuGly 127
Db      509 -----GATGAAGGATTTTATCATGTCGTCTGGCAATGAACAAATATGGA 550

```

```

Qy      128 ValValAlaSerGlnValAlaValValLysLeuAlaThrLeuGluAspPheSerLeuHis 147
Db      551 GCCATTCTTAGTCAAAAAGCTCATCTGCTTATCACTATTCTTCGATTTCGAAGTCCAG 610
Qy      148 ProGluSerGlnIleValGluGluAsnGlyThrAlaArgPheGluCysHisThrLysGly 167
Db      611 CCNAATTTCCACTGAGTCCACGAAGTGGAGTTGCTCGATTTCGATGCAAGATTTTCATCC 670
Qy      168 LeuProAlaProIleIleThrTrpGluLysAspGlnValThrValPro---GluGluPro 186
Db      671 CACCTCTCTGCAGTCTAACATGGAGTTCATCGACACACTCTTACCTATGACTATGGAC 730
Qy      187 ArgLeuIleThrLeuProLysTrpLeuLeuGlnIleLeuAspValGlnAspSerAspAla 206
Db      731 AGGATAACTGCCCTACCAACAGGAGTATTGCAGATCTATGATGTTCAGCCAAAGGATTTCT 790
Qy      207 GlySerThrArgCysValAlaThrAsnSerAlaArgGlnArgPheSerGlnGluAlaSer 226
Db      791 GGAAATTATCGTTGTTATTTGCTGCCACTGTAGCCACCGACGCTAAAGATATGGAGGCTCG 850
Qy      227 LeuThrValAlaLeuArgGlySerLeuGluAlaThrArgGlyGlnAspValIleVal 246
Db      851 CTRACTGTGAT-----CCAGCTAAGGAGTCAAAATCCTCCACACACCAACCAATATA 904
Qy      247 AlaAlaProGluAsnThrThrValValSerGlyGlnAsnValValMetGluCysValAla 266
Db      905 GCAGGTCCACAGAACATAACAACATCTCTTCATCAGACTGTAGTTTGGATGTCATGGCC 964
Qy      267 SerAlaAspProThrProPheValSerTrpValArgGlnAspGlyLysProIleSerThr 286
Db      965 ACAGGAATTCGCAACCAATCATTTCTTGGAGCGCTTGATCACAATATCCATTCATGTC 1024
Qy      287 ---AspValIleValLeuGlyArgThrAsnLeuLeuIleAlaSerAlaGlnProArgHis 305
Db      1025 TTTAATCTCGGTACTTGGAAATGTAATCTCATGATATCTCATGTGAGGTACCAACAT 1084
Qy      306 SerGlyValThrValCysArgAlaAsnLysProLeuThrArgAspPheAlaThrAlaAla 325
Db      1085 GCTGGAGTATATGTTTGTGGGCCACTACCCCTGGCCACACGCAACTTTACAGTTGCTATG 1144
Qy      326 AlaGluLeuArgValLeuAlaAlaProAlaIleSerGlnAlaProGluAlaLeuSerArg 345
Db      1145 GCAACTTTAACTGTATTAGTCTCTCTTCATTTGTGTAATGGCCAGAAAGTTTAAACAGG 1204
Qy      346 ThrArgAlaSerThrAlaArgPheValCysArgAlaSerGlyGluProArgProAlaLeu 365
Db      1205 CCTCGAGCTGCACCTGCTCGATTTGTGTGAGGAGGAAGAAATCCCTCTCCCAAGATG 1264
Qy      366 HisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgValLysValGlnGlyGly 385
Db      1265 TCATGTTGAAAATGGAGGAGNAGATACATTCGATGTTAGATTTAAATGTAC----- 1318
Qy      386 GlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyThrGlnCysVal 405
Db      1319 AACAGTAAATGGTAATTAACCCAGATTATCTCGAAGATGATGCTATTATTCAGTGCATG 1378
Qy      406 AlaGluAsnSerAlaGlyThrAlaCysAlaAlaProLeuAlaValValValArgGlu 425
Db      1379 GCTGAGATAGCCAAAGGATCTATTATATCTAGAGCCAGACTGATGTATGTATGTCAGAA 1438
Qy      426 GlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerSerValLeu 445
Db      1439 GACAGACCCAGTCTCTCTATATGATGCTGTGAACCATGTCAGCTCAGCTCAGCCATCTT 1498
Qy      446 ValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisThr 465
Db      1499 TTAGCTGGAGAGGCCACTTTATAATTTACAGAAAGTCATTGCTCTATTCTGTACACTAC 1558
Qy      466 GlnLysAlaArgGlyValAspAsnValGluThrGlnPheAlaValAspAsnAspThrThr 485
Db      1559 ATGAAAGCAGAGGTTTAAATAATGAAGATATCAAGTAGTATCATCGGAAATGACACAACT 1618

```

```

Qy 486 GluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPheTyrValValAlaTyr 505
Db 1619 CATTATATTATGATGACTAGAGCCTGCCAGCAATTATATCTTCTACATGTGACATAT 1678
Qy 506 SerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThrLeuAspAspVal 525
Db 1679 ATGCCAATGGGAGCGAGCAGATCTCTGACCATGTGACACAGAAATATCTCTAGAGATGTT 1738
Qy 526 ProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAspIleArgValAla 545
Db 1739 CCCCTGAGACCTCTCGAAATAGTTTGACAAGTCGAAGTCCACCTGATATCTCATCTCC 1798
Qy 546 TrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLysTyrLysIleGluTyr 565
Db 1799 TGGCTGCCAATCCAGCCCAATATCGCGCGGCCAAGTGTGTGTATCGCTTCTCTTTC 1858
Qy 566 GlyLeuGlyLysGluAspGlnValPheSerThrGluValProGlyAsnGluThrGlnLeu 585
Db 1859 CGCCTAAGTACTGGAATTCATCAAGTCTCGAGCTCCCGGGGACACGCGATGATAC 1918
Qy 586 ThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgIleSerAlaGlyThrGly 605
Db 1919 CTTTGGNAGCCTGAACTGACAGTGTCTACCTGTGTGCGATATCTGTGCCACACAGA 1978
Qy 606 AlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrProGlyValHisAsnGln 625
Db 1979 GTGGGCTGGGAGAGTCATGATGACTTCATAGGACGCCCAAGCT--ACAAGC 2035
Qy 626 SerHisValProPheAlaProAlaGluLeuLysValArg---AlaLysMetGluSerLeu 644
Db 2036 GTGAAAGCCCCTAAGTCTCCA--GAGTTGATTTGGAGCCTCTGAACTGTACCAACCAT 2092
Qy 645 ValValSerTrpGlnProProHisProThr--GlnIleSerGlyTyrLysLeuTyr 663
Db 2093 TCTGTAGGTGGGAGCAAGATGTAGAGGACACAGCTGCTATTTCAGGCTACAGCTGTAC 2152
Qy 664 TrpGlyGluValGlyThrGluGluGluAlaAspGlyAspArgProProGlyGlyArgGly 683
Db 2153 TACAAGGAAGAAGGCGACGAGAAAT----- 2179
Qy 684 AspGlnAlaTrpAspValGlyProValArgLeuLysLysLysValLysGlnTyrGluLeu 703
Db 2180 -----GGGCCCATTTTCTTGATACCAAGGACCTACTCTATACTCTC 2221
Qy 704 ThrGlnLeuValProGlyArgProTyrGluValLysLeuValAlaPheAsnLysHisGlu 723
Db 2222 AGTGGCTTAGCCACAGAAATATCATGTGAGACTCTCTGGCTTACACACATAGAC 2281
Qy 724 AspGlyTyrAlaAlaValTrpLysGlyThrGluLysAlaProThrProAspLeuPro 743
Db 2282 GATGGCTATCAGGCA-----GATCAGACTGTGACACTCCAGGATGCGTG 2326
Qy 744 IleGlnArgGly-----ProLeuProProAlaHisValHisAlaGluSer 759
Db 2327 TCTGTTCGTGATGCGCATGCTCCCTCTCCACACACCCACCATCTCTATGCGAAGCT 2386
Qy 760 AsnSerSerThrSerIleTrpLeuArgTrpLysLysProAspPheThrThrValLysIle 779
Db 2387 AACCTCATCTTCATCTCTGCACTGGAGGCGCTGCATTCACCGCTGCAAAATC 2446
Qy 780 ValAsnTyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTyr 799
Db 2447 ATTAACATACCATCCGCTGAATCTCTGTGGCTGCGAGATGCTCTTGTGTTCTGTAC 2506
Qy 800 TyrThrSerSerGlyGluAspIleLeuIleGlyLysLysProPheThrLysTyrGlu 819
Db 2507 CTTCAAAACATCAGAACTCAGATGTTGTTTCAAGGTCTAGAACCAACCAACCAATACGAA 2566
Qy 820 PheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArg 839
Db 2567 TTTGCCGTTCGATTACATGTGGATCAGCTTTTCCAGTCTCTGGAGCCCTGTAGTCTACCAT 2626
Qy 840 SerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThr

```

```

Db 2627 TCTACTCTTCCAGAACACGACGAGCCACCACCTGGAGTAAAGTGACATTAATAGAG 2686
Qy 860 ProSerThrValArgLeuHisTrpCysProThrThrGluProAsnGlyGluIleValGlu 879
Db 2687 GATGACACTGCCCTGTTTCTTGGAAACCCCTGTATGGCCAGAAACAGTTGTGACCCGC 2746
Qy 880 TyrLeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThr 899
Db 2747 TATACTATCTTATGATCATCTAGAAAGCGCTGATTCGACGAGAGTGGCAGGTCTTACAC 2806
Qy 900 ThrGluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhe 919
Db 2807 CGTGAAGGGGCAATACCATGGCTTTGTAGAAAACTTGTAGCAGAAATGTGTACATT 2866
Qy 920 PheLysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspVal 939
Db 2867 GTCAGATATCTGCATCCATAGAGTGGGAGGAGGACCCCTTTCAATTCCTGTGGAGCTG 2926
Qy 940 IleThrLeuGlnGluThrPheSerAspSer----- 949
Db 2927 GCAGTACTTCCAAAGGAAACCTCTGAATCAATCAGAGGCCCAAGCGTTTAGATTCTGCT 2986
Qy 950 -----LeuaspValHisAlaValThrGlyIleIle 959
Db 2987 GATGCCAAAGTTTATTCAGGATATTACCATCTGGACCCAAAAATCAATGATGGCATTGT 3046
Qy 960 ValGlyValCysLeuGlyLeuLeuCysAlaLeuAlaCysMetCysAlaGlyLeuArgGln 979
Db 3047 GTAGGTGTGGCATAGCTTGACCTGCATCTCATCTGTGTCTCATCTCTGTATATACCGA 3106
Qy 980 SerSerHisArgGluAla-----LeuPro 987
Db 3107 AGTAAAGCCAGGAAATCATCTGCTTCCAGACGGCACAGAACTGAACTCAACAGTTACCT 3166
Qy 988 GlyLeuSerSerSerGlyThrProGlyAsn 997
Db 3167 CGTACCAGTGCCTCTTAGCTAGTGGAAAT 3196

RESULT 4
US-09-877-730-11
; Sequence 11, Application US/09877730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730.
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 2976
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-11

Alignment Scores:
Pred. No.: 1,79e-102 Length: 2976
Score: 1691.50 Matches: 375
Percent Similarity: 55.79% Conservative: 160
Best Local Similarity: 39.10% Mismatches: 347
Query Match: 25.61% Indels: 77
DB: 4 Gaps: 15

```



```
QY 791 LeuArgAsnAlaSerLeuValThrTyrThrSerSerGlyGluAspIleLeuLeuGly 810
Db 2056 CTGCAGATGCTCTTGGTCTGTACCTTCAACATCAGAACTCACATGTTGGTTCAA 2115
QY 811 GlyLeuLysProPheThrLysTyrGluPheAlaValGlnSerHisGlyValAspMetAsp 830
Db 2116 GGTCTAGAACCAACACCAATGCGTTGCGATTACATGTGGATCGCTTTCC 2175
QY 831 GlyProPheGlySerValValGluArgSerThrLeuProAspArgProSerThrProPro 850
Db 2176 AGTCCTTGGAGCCTGTAGTCTACCAATTTCTCTTCCAGAGCACAGAGGCCCAACA 2235
QY 851 SerAspLeuArgLeuProLeuThrProSerThrValArgLeuHisTrpCysProPro 870
Db 2236 GTTGGAGTAAAGTGACATTAATAGAGGATGACACTGCTGTTCTTGGAAACCCCT 2295
QY 871 ThrGluProAsnGlyGluIleValGluTyrLeuIleLeuLysSerAsnAsnHisThrGln 890
Db 2296 GATGGCCCAAGAACAGTGTGACCGCTATCTATCTATATGCTATGATCATCTAGGAGGCTGG 2355
QY 891 ProGluHisGlnTrpThrLeuLeuThrThrGluGlyAsnIlePheSerAlaGluValHis 910
Db 2356 ATTGCAGGAGGTGGCGAGTCTTACACCGTGAAGGGCGCAATAACCATGGCTTGTGTAGAA 2415
QY 911 GlyLeuGluSerAspThrArgTyrPhePheLeuMetGlyAlaArgThrGluValGlyPro 930
Db 2416 AACTTGGTAGCAGGAATGTGTACATGTGTCAAGATATCTCATCCATGAGGTGGAGAA 2475
QY 931 GlyProPheSerArgLeuGlnAspValIleThrLeuGlnGluThrPheSerAspSer--- 949
Db 2476 GGACCCCTTTTCAAAATCTGTGGAGTGGCAGTACTTCCAAAGGAAACCTCTGAATCAAAAT 2535
QY 950 -----Leu 950
Db 2536 CAGAGGCCCAAGCGTTTAGAATCTGCTGATGCCAAAGTTTATTACGATATTACCATCTG 2595
QY 951 AspValHisAlaValThrGlyIleIleValGlyValCysLeuGlyLeuLeuCysLeuLeu 970
Db 2596 GACCAAAATCAATGACTGCTGATGTGTAGTGTGGCATGTGCTTGCATCTGCTGCTCTC 2655
QY 971 AlaCysMetCysAlaGlyLeuArgGlnSerSerHisArgGluAla----- 985
Db 2656 ATCTGTGTTCTCATCTTGATATACCGAAGTAAAGCCAGGAAATCATCTGCTTCCAAAGAC 2715
QY 986 -----LeuProGlyLeuSerSerSerGlyThrProGlyAsn 997
Db 2716 GCACAGAAATGGAATCAACAGTACCTCTGATCCAGTCCGTCAGTCCCTTAGCTAGTGAAT 2772
RESULT 5
; Sequence 17, Application US/09877730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding the
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3219
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-17
```

Alignment Scores:

Pred. No.: 2e-102 Length: 3219
Score: 1691.50 Matches: 375
Percent Similarity: 55.79% Conservative: 160
Best Local Similarity: 39.10% Mismatches: 347
Query Match: 25.61% Indels: 77
DB: 4 Gaps: 15

US-09-754-997A-2 (1-1252) x US-09-877-730-17 (1-3219)

```
QY 79 GluHisGluAsnLeuHisLeuLeuProAsnGlySerLeuTrpLeuSerSerProLeuGlu 98
Db 7 GAAATAAAGCATCGATGAGTCTCTTTCTAACGCGCTCTTTATACATCAGT-----GAG 57
QY 99 GlnGluAspSerAspGluAlaLeuArgIleTrpLysValThrGluGlySerTyr 118
Db 58 GTGGAAGGCGAGGCGAGGAGCAGTCC-----GATGAAGGATTTTAT 99
QY 119 SerCysLeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValValLysLeu 138
Db 100 CAGTCTTGGCAATGACAAATATATGAGCCATTTCTTAGTCAAAAAGCTCATCTTCCCTTA 159
QY 139 AlaThrLeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThr 158
Db 160 TCAACTATTCTGCAATTTGAAAGTCCAGGCAATTTCCACTGAGGTCCACGAGGTGGAGTT 219
QY 159 AlaArgPheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAsp 178
Db 220 GCTCGATTGTCATGCAAGATTTATCCCAACCTCTCTGAGTCATATCAATGGAGTTTCAAT 279
QY 179 GlnValThrValPro---GluGluProArgLeuIleThrLeuProLysTrpLeuGln 197
Db 280 CGGACACTCTACTATGATGATGACGAGGATAATCGCTTACCAACAGGAGTATTGCGAG 339
QY 198 IleLeuAspValGlnAspSerAspAlaGlySerTyrArgCysValAlaThrAsnSerAla 217
Db 340 ATCTATGATGTGAGCAAGGATTTCTGGAATATATCGTTGATTATGCTGCCACTGTAGCC 399
QY 218 ArgGlnArgPheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGluAla 237
Db 400 CACCGACGTAAAGATGATGAGGCGCTGCTTAATCTGTGATT-----CCAGCTAAGGAGTCA 453
QY 238 ThrArgGlyGlnAspValIleValAlaAlaProGluAsnThrThrValValSerGly 257
Db 454 AAATCTCTCCACACACACCAATTTATAGCAGTCCACAGAACATCAACAATCTCTTCAT 513
QY 258 GlnAsnValValMetGluCysValAlaSerAlaAspProThrProPheValSerTrpVal 277
Db 514 CAGACTGTAGTTTGGATGTCATGCGCACGGAATCCCAAAACCAATCATCTTCTTGGAGC 573
QY 278 ArgGlnAspGlyLysProIleSerThr---AspValIleValLeuGlyArgThrAsnLeu 296
Db 574 CGCTTGTATCACAATCTCATGATGTCTTTAATATCTCGGGTACTTGGAAATGGTAATCTC 633
QY 297 LeuIleAlaSerAlaGlnProArgHisSerGlyValTyrValCysArgAlaAsnLysPro 316
Db 634 ATGATATCTGATGTGAGGTACACATGCTGGAGTATATGTTGTGGGCCCATCTCCCT 693
QY 317 LeuThrArgAspPheAlaThrAlaAlaGluLeuArgValLeuAlaAlaProAlaIle 336
Db 694 GGCACACGCAACTTTACAGTTGCTATGCGCAACTTTAACTGTATTAGTCTCTCTCTCATTT 753
QY 337 SerGlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArg 356
Db 754 GTTGAATGGCCAGAAAGTTTAAACAGGCTCGAGCTGGCAGTCTCGATGTTGTGTGTCAG 813
QY 357 AlaSerGlyGluProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgPro 376
Db 814 GCAGAAGGAATCCCTCTCCCAAGATGTCTATGTTGAAAATGGAAGGAGATATCATTCG 873
QY 377 AsnGlyArgValLysValGlnGlyGlyGlyGlySerLeuValIleThrGlnIleGlyLeu 396
Db 377 AsnGlyArgValLysValGlnGlyGlyGlyGlySerLeuValIleThrGlnIleGlyLeu 396
```

874 AATGGTAGAATTAAATGTAC-----AACAGTAAATTCGTAAATTAACACAGATTATTCCT 927
Db
397 GlnAspAlaGlyTyrTyrGlnCysValAlaGluAsnSerAlaGlyThrAlaCysAlaAla 416
Qy
928 GAAGATGATGCTATTATTCAGTGCATGCTGAGATAGCAAGGATCTATTTTATCTAGA 987
Db
417 AlaProLeuAlaValValAlaGluGlyLeuProSerAlaProThrArgValThrAla 436
Qy
988 GCCAGACTGACTGTAGTGTGTCAGAAAGACAGACAGACCCAGTGCCTCCCTATATATGTATGCT 1047
Db
437 ThrProLeuSerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGlu 456
Qy
1048 GAAACCATGTCAAGCTCAGCCATTCTTTAGCTGGGAGAGGCCACTTTTAAATTAATTCAGAC 1107
Db
457 GlnIleGlyPheSerLeuHisTyrGlnLysAlaAaGlyValAspAsnValGluTyr 476
Qy
1108 AAGATCATTCCTTCTGACATACATGAAAGCAGAGGTTTAAATAATCAAGATAT 1167
Db
477 GlnPheAlaValAsnAsnAspThrThrGluLeuGlnValArgAspLeuGluProAsnThr 496
Qy
1168 CAAGTAGTCTCGGAATGACCAACTCATTTATTTATTTGATGACTTAGAGCTGCCAGC 1227
Db
497 AspTyrGluPheTyrValValAlaTyrSerGlnLeuGlyValAspSerArgThrSerPro 516
Qy
1228 AATTATACCTTCTACATTTGATATATGCAATGGGAGCCAGCCAGATGCTGACCAT 1287
Db
517 AlaLeuValHisThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSer 536
Qy
1288 GTGACACAGAACTACTCAGAGAGATGTTCCCTGAGACCTCTCGAAATTAATTTGACAGT 1347
Db
537 ProAsnProSerAspIleArgValAlaTrpLeuProLeuProSerSerLeuSerAsnGly 556
Qy
1348 CGAAGTCCCACTGATATTCCTATCTCTGCTGCCAATCCAGCCCAATATCGCGGGGC 1407
Db
557 GlnValLeuTyrTyrLysIleGluTyrGlyLeuGlyLysGluAspGlnValPheSerThr 576
Qy
1408 CAAGTGGTCTGTATGCTGTGCTTTCCGCCTAAGTACTGAGAAATTCATCAAGTTCTG 1467
Db
577 GluValProGlyAsnGluThrGlnLeuThrLeuAsnSerLeuGlnProAsnLysValTyr 596
Qy
1468 GAGCTCCCGGAGCCACCGATGATGATCTTTTGGAGGCTGAAACCTGACAGATGCTTAC 1527
Db
597 ArgValArgIleSerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGln 616
Qy
1528 CTGGTTCGATTTACTGCTGCCACAGAGTGGGCTGGGAGATCATCATGAGATTCA 1587
Db
617 HisArgThrProGlyValHisGlnSerHisValProPheAlaProAlaGluLeuLys 636
Qy
1588 CATAGACGCCCAAGCT--ACAAGCGTGAAGCCCTTAAGTCTCCA--GAGTTGCAAT 1641
Db
637 ValArg--AlaLysMetGluSerLeuValValSerTrpGlnProProHisProThr 655
Qy
1642 TTGGAGCCTCTGAACTGTATACCATTTCTGTGAGGTGCAGCAGATGTAGAGGACACA 1701
Db
656 ---GlnIleSerGlyTyrLysLeuTyrTrpGlyGluValGlyThrGluGluAlaAsp 674
Qy
1702 GCTGCTATTTCAGGCTACAGCTGTACTACAAGGAAGAGGGCAGCAGAGAAT----- 1755
Db
675 GlyAspArgProGlyGlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeu 694
Qy
1756 -----GGGCCCAATTTTCTTG 1770
Db
695 LysLysValValLysGlnTyrGluLeuThrGlnLeuValProGlyArgProTyrGluVal 714
Qy
1771 GATACCAAGGACTCTCTATCTCTCAGTGGCTTAGACCCAGAGAAATATCATGTG 1830
Db
715 LysLeuValAlaPheAsnLysHisGluAspGlyTyrAlaAlaValTrpLysGlyLysThr 734
Qy
1831 AGACTCTGCTTACCAACATAGAGATGGCTATCAGCA-----GAT 1875
Db
735 GluLysAlaProThrProAspLeuProIleGlnArgGly-----ProProLeu 750
Qy
1876 CAGACTGTCCAGCACTCCAGATGCGGTCTGTCTGTGATCGCATGTGCTCCTCCTCCACCA 1935
Db

751 ProProAlaHisValHisAlaGluSerAsnSerThrSerIleTrpLeuArgTrpLys 770
Qy
1936 CCACCCCACTCTCTATGCGAAGCTAAACCTCATCTTCCATCTTCTGCTGACCTGGAG 1995
Db
771 LysProAspPheThrThrValLysIleValAsnTyrThrValArgPheGlyProTrpGly 790
Qy
1996 AGGCTGATCTACCGCTGCAAAATCAATTAACATACACCATCCGCTGTAATCTCTGTGGC 2055
Db
791 LeuArgAsnAlaSerLeuValThrTyrTyrThrSerSerGlyGluAspIleLeuIleGly 810
Qy
2056 CTGCAAGATGTTCTTTGGTCTGTACCTTCAACATCAGAAATCAATCATGTTGTTCAA 2115
Db
811 GlyLeuLysProPheThrLysTyrGluPheAlaValGlnSerHisGlyValAspMetAsp 830
Qy
2116 GGTCTAGAACCAACACCAATACGAATTTGCGGTTCGATTACATGTGATCAGCTTCC 2175
Db
831 GlyProPheGlySerValValGluArgSerThrLeuProAspArgProSerThrProPro 850
Qy
2176 AGTCCCTGGAGCCCTGTAGTCTACCATTTCTTCCAGAAGCCAGCAGGCCACCA 2235
Db
851 SerAspLeuArgLeuSerProLeuThrProSerThrValArgLeuHisTyrCysProPro 870
Qy
2236 GTTGAGTAAAGTGACATTAATAGAGATGACACTGCCCTGGTTCTTGGAAACCCCT 2295
Db
871 ThrGluProAsnGlyGluIleValGluTyrLeuIleLeuTyrSerAsnAsnHisThrGln 890
Qy
2296 GATGCCCCAGAAACAGTTGTGACCGCTATACTATCTTATATGTCATCTAGGAAGCCTGG 2355
Db
891 ProGluHisGlnTyrThrLeuLeuThrThrGluGlyAsnIlePheSerAlaGluValHis 910
Qy
2356 ATTGAGAGAGTGGCAGTCTTACACCGTGAAGGGGCAATAACCATGCTTTGCTAGAA 2415
Db
911 GlyLeuGluSerAspThrArgTyrPhePheLysMetGlyAlaArgThrGluValGlyPro 930
Qy
2416 AACTGTGAGAGGAATGTGTACATTCATGATATTCGATCCATCCATGAGGTGGAGAA 2475
Db
931 GlyProPheSerArgLeuGlnAspValIleThrLeuGlnGluThrPheSerAspSer--- 949
Qy
2476 GGACCTTTTCAAAATCTGTGGAGTGGCAGTACTTCCAAAGGAACCTCTGAATCAAT 2535
Db
950 -----Leu 950
Qy
2536 CAGAGCCCAAGCGTTAGATTCTGCTGATGCCAAAGTTTATTTCAGGATATTTACCATCTG 2595
Db
951 AspValHisAlaValThrGlyIleValGlyValCysLeuGlyLeuLeuCysLeuLeu 970
Qy
2596 GACCAAAATCAATGACTGGCATTTGCTGTAGTGTGGCATAGCCTTGACCTGCATCTC 2655
Db
971 AlaCysMetCysAlaGlyLeuArgGlnSerSerHisArgGluAla----- 985
Qy
2656 ATCTGTGTCTCTCATCTTGATATACCGAAGTAAAGCCAGGAATCATCTGCTTCCAGAGC 2715
Db
986 -----LeuProGlyLeuSerSerSerGlyThrProGlyAsn 997
Qy
2716 GCACAGATGGAACTCAACAGTTACTCTGACAGTGGCTCTCTTAGCTAGTGGAAAT 2772
Db

RESULT 6

US-09-877-730-5

; Sequence 5, Application US/09877730

; Patent No. 6465632

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Scoville, John

; APPLICANT: Turner, C. Alexander Jr.

; APPLICANT: Friedrich, Glenn

; APPLICANT: Abuin, Alejandro

; APPLICANT: Zambrowicz, Brian

; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding t

; FILE REFERENCE: LEX-0189-USA

; CURRENT APPLICATION NUMBER: US/09/877,730

; CURRENT FILING DATE: 2001-06-08

Db 1554 ----- 1554
Qy 646 ValSerTrpGlnProProHisProThrGlnIleSerGlyTyrLysLeuTyrTrpGly 665
Db 1554 ----- 1554
Qy 666 GluValGlyThrGluGluGluAlaAspGlyAspArgProGlyGlyArgGlyAspGln 685
Db 1554 ----- 1554
Qy 686 AlaTrpAspValGlyProValArgLeuLysLysValLysGlnTyrGluLeuThrGln 705
Db 1554 ----- 1554
Qy 706 LeuValProGlyArgProTyrGluValLysLeuValAlaPheAsnLysHisGluAspGly 725
Db 1555 ----- 1555
Qy 726 TyrAlaAlaValTrpLysGlyThrGluLysAlaProThrProAspLeuProIleGln 745
Db 1603 TATCAGGCA-----GATCAGACTGTGCAGACTCCAGGATGGTGTCTGTT 1647
Qy 746 ArgGly-----ProLeuProProAlaHisValHisAlaGluSerAsnSer 761
Db 1648 CGTGATCCAGTGGTCCCTCCACCCACCCACCATCTCTATGCGAGGCTTAACACC 1707
Qy 762 SerThrSerIleTrpLeuArgTrpLysLysProAspPheThrThrValLysIleValAsn 781
Db 1708 TCATCTTCATCTCTCTGATCTGGAGGAGGCTGCATTCACCGTGCACAAATCATTAAC 1767
Qy 782 TyrThrValArgPheGlyProTrpGlyLeuArgAlaSerLeuValThrTyrTrpThr 801
Db 1768 TACACCATCCGCTGTAATCCCTGGCTGAGGATGCTCTCTGTTCTGTACCTCAA 1827
Qy 802 SerSerGlyGluAspIleLeuIleGlyLysLysProPheThrThrLysTrpGluPheAla 821
Db 1828 ACATCAGAACTCATCTGGTTCAGGCTTAGAACCAACCAACCAATACGAAATTTGCC 1887
Qy 822 ValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSerThr 841
Db 1888 GTTCGATTACATGTGATCAGCTTCCAGTCTCTGGAGCCCTGTAGTCTACCATCTACT 1947
Qy 842 LeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrProSer 861
Db 1948 CTTCCAGAACACACAGGCCCACAGTGGAGTAAAGTGACATTAAATAGAGGATGAC 2007
Qy 862 ThrValArgLeuHisTrpCypProThrGluProAsnGlyGluIleValGluTyrLeu 881
Db 2008 ACTGCCCTGGTTCTTGAAACCCCTGATGCGCCAGAAACAGTTGTGACCCGCTATACT 2067
Qy 882 IleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThrGlu 901
Db 2068 ATCTTATATGATCTAGGAGCCCTGGATTGAGAGAGTGGAGGCTTTACACCGTGA 2127
Qy 902 GlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePheLys 921
Db 2128 GGGGCAATAACCATGGCTTCTAGAAACCTTGGTAGCAGGAAATGTATCATCTCAAG 2187
Qy 922 MetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIleThr 941
Db 2188 ATATCTGCATCAATGAGTGGAGAGGAGCCCTTTTCAAAATCTCTGTGGAGCTGCAGTA 2247
Qy 942 LeuGlnGluThrPheSerAspSer----- 949
Db 2248 CTTCCAAAGAAACCTCTGAAATCAAAATCAGAGGCCCAAGCGTTTATGATCTGTGATGCC 2307
Qy 950 -----LeuAspValHisAlaValThrGlyIleIleValGly 961
Db 2308 AAGTTTATTCCAGGATATTAATCTGACCAAAATCAATGACTGGCAATGTCTGTAGGT 2367
Qy 962 ValCysLeuGlyLeuLeuCysLeuAlaCysMetCysAlaGlyLeuArgGlnSerSer 981
Db 2368 GTTGCCATAGGCTTGACCTGCATCTCTGTGTCTCATCTTGTATATACCGAAGTAA 2427

Qy 982 HisArgGluAla-----LeuProGlyLeu 989
Db 2428 GCCAGGAATCATCTGCTTCCAGACGCGCAGAGTGAACCTCAACAGTTTACCTCGTACC 2487
Qy 990 SerSerSerGlyThrProGlyAsn 997
Db 2488 AGTCCCTCTCTAGCTAGTGGAAAT 2511

RESULT 7

US-09-877-730-9
; Sequence 9, Application US/09877730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2958
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-9

Alignment Scores:
Pred. No.: 3,56e-82 Length: 2958
Score: 1384.00 Matches: 334
Percent Similarity: 46.11% Conservative: 140
Best Local Similarity: 32.49% Mismatches: 310
Query Match: 20.96% Indels: 244
DB: 4 Gaps: 14

US-09-754-997A-2 (1-1252) x US-09-877-730-9 (1-2958)

Qy 9 GlyLeuLeuValLeuThrPheCysLeuSerAlaArgGlyGluLeuProLeuProGln 28
Db 43 GGGATGCTGCTCGCGCGCTCTCTGCTGCTGCTCAGT-----CCTTTGCCAGA 96
Qy 29 GluThrThrVal---LysLeuSerCysAspGluGlyProLeuGlnValIleLeuGlyPro 47
Db 97 GTGTGGTGTCTTACGCAACTGCTTTTGTAAAGAACCAACAGGATGTAACGTGCACAAGA 156
Qy 48 GluGlnAlaValValLeuAspCysThrLeuGlyAlaThrAlaAlaGlyProProThrArg 67
Db 157 AAGGACCCAGTGTGTATTGATGGCAGCTCAGGAGAGTT-----CCTATTAAAG 207
Qy 68 ValThrTrpSerLysAspGlyAspThrValLeuGluHisGluAsnLeuHisLeuLeuPro 87
Db 208 GTCCATGTTGTAATAATGGAGCAAAATGCTGAAAATAAACGGATCGAGGTTCTTTCT 267
Qy 88 AsnGlySerLeuTrpLeuSerSerProLeuGluGlnGluAspSerAspGluGluAla 107
Db 268 AACGGCTCTTTATACATCAGT-----GAGGTGGAGGCGAGGAGGAGCAGTCC 318
Qy 108 LeuArgIleTrpLysValThrGluGlySerTyrSerCysLeuAlaHisSerProLeuGly 127
Db 319 -----GATGAAGGATTTTATCAGTCTTGGCAATGCAACAAATATGGA 360
Qy 128 ValValAlaSerGlnValAlaValValLysLeuAlaThrLeuGluAspPheSerLeuHis 147
Db 361 GCCATTCTTAGTCAAAAGGCTCATCTTGCCTTATCACTATTCTTCGCAATTTGAAGTCCAG 420


```
Qy 457 GlnIleIleGlyPheSerLeuHisTyrGlnIysAlaArgGlyValAspAenValGluTyr 476
Db 1108 AAGTCATTGCCATTCTGTACACTATACGAAGCAGAGGTTTAAATAATGAAGAGTAT 1167
Qy 477 GlnPheAlaValAenAenAspThrThrGluLeuGlnValArgAspLeuGluProAsnThr 496
Db 1168 CAAGTAGTCATCGGAATATGACACAACTCATATATATTATGATGACTTAGAGCCGCCAGC 1227
Qy 497 AspTyrGluPheTyrValAlaTyrSerGlnLeuGlyAlaSerArgThrSerPro 516
Db 1228 AATTATATCTTCTACATTGTAGCATATATGCAATGGAGCCAGCCAGATGTCTGACCAT 1287
Qy 517 AlaLeuValHisThrLeuAspValProSerAlaAlaProGlnLeuThrLeuSerSer 536
Db 1288 GTGACACAGAACTCTAGAGGATGACCCGAGA----- 1320
Qy 537 ProAenProSerAspIleAlaGValAlaTriLeuProLeuProSerSerLeuSerAsnGly 556
Db 1320 ----- 1320
Qy 557 GlnValLeuLysTyrLysIleGluTyrGlyLeuGlyLysGluAspGlnValPheSerThr 576
Db 1320 ----- 1320
Qy 577 GluValProGlyAenGluThrGlnLeuThrLeuAenSerLeuGlnProAenLysValTyr 596
Db 1320 ----- 1320
Qy 597 ArgValArgIleSerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGln 616
Db 1320 ----- 1320
Qy 617 HisArgThrProGlyValHisAenGlnSerHisValProPheAlaProAlaGluLeuLys 636
Db 1320 ----- 1320
Qy 637 ValArgAlaLysMetGluSerLeuValValSerTrpGlnProProHisProThrGln 656
Db 1320 ----- 1320
Qy 657 IleSerGlyTyrLysLeuTyrTrpGlyGluValGlyThrGluGluGluAlaAspGlyAsp 676
Db 1320 ----- 1320
Qy 677 ArgProProGlyGlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLysLys 696
Db 1320 ----- 1320
Qy 697 LysValLysGlnTyrGluLeuThrGlnLeuValProGlyArgProTyrGluValLysLeu 716
Db 1321 -----AGAAAATATCATGTGAGACTC 1341
Qy 717 ValAlaPheAenLysHisGluAspGlyTyrAlaAlaValTrpLysGlyLysThrGluLys 736
Db 1342 CTGGCTTACAACAACATAGACATGGCTATCAGGCA-----GATCAGACT 1386
Qy 737 AlaProProAspLeuProIleGlnArgGly-----ProProLeuProPro 752
Db 1387 GTCAGCACTCCAGGATGGTGTCTGTCGTGATCGATGGTCCCTCCACCCACCCACC 1446
Qy 753 AlaHisValHisAlaGluSerAenSerSerThrSerIleTrpLeuArgTrpLysLysPro 772
Db 1447 CACCATCTCTATGCGAAGGCTTAACACCTCATCTTCATCTTCCCTGCACTGGAGGGGCT 1506
Qy 773 AspPheThrValLysIleValAenTyrThrValArgPheGlyProTrpGlyLeuArg 792
Db 1507 GCATTACCGCTGCACAATCATTTAATACCATCCGCTGTATCTCTGTGGCTGAG 1566
Qy 793 AsnAlaSerLeuValThrTyrThrSerSerGlyGluAspIleLeuIleGlyGlyLeu 812
Db 1567 AATGCTCTTTGGTTCTGTACCTTCAACATCAGAAACTCACATGTTGTTCAAGGTCTA 1626
Qy 813 LysProPheThrLysTyrGluPheAlaValGlnSerHisGlyValAspMetAspGlyPro 832
```

```
Db 1627 GAACCAAAACCAACCAATACGAATTTGCGTTCGATTACATGTGGATCAGCTTTCCAGTCT 1686
Qy 833 PheGlySerValValGluArgSerThrLeuProAspArgProSerThrProProSerAsp 852
Db 1687 TGGAGCCCTGTAGTCTACCATTTCTACTTCTCCAGAAGCACACAGCAGGCCCCACCAAGTTGGA 1746
Qy 853 LeuArgLeuSerProLeuThrProSerThrValArgLeuHisTyrCyseProThrGlu 872
Db 1747 GTAAAGTAGTACCAATTAATAGAGTAGACACTGCCCTGGTTTCTTGGAAACCCCTGATGCG 1806
Qy 873 ProAenGlyGluIleValGluTyrLeuIleLeuTyrSerAsnAenHisThrGlnProGlu 892
Db 1807 CCAGAAACAGTTGTGACCGCTATATCTTATATGTCATCTAGGAAGSCCTGGATTGCA 1866
Qy 893 HisGlnTrpThrLeuLeuThrThrGluGlyAenIlePheSerAlaGluValHisGlyLeu 912
Db 1867 GGAGAGTGGCAGTCTTACACCGTGAAGGGGCAATAACCATGCTTGTCTAGAAAACCTTG 1926
Qy 913 GluSerAspThrArgTyrPhePheLysMetGlyAlaArgThrGluValGlyProGlyPro 932
Db 1927 GTAGCAGGAATAATGTGTACATTTGTCAAGATATCTGCATCCAAATGAGTGGGAGAGGCC 1986
Qy 933 PheSerArgLeuGlnAspValIleThrLeuGlnGluThrPheSerAspSer----- 949
Db 1987 TTTTCAATCTGTGGAGCTGGCAGTACTTCCAAAGGAAACCTCTGAATCAATCAGAGS 2046
Qy 950 -----LeuAspVal 952
Db 2047 CCCAAGCGTTAGATTCTGCTGATGCCAAAGTTTATTTCAGGATATTACCATCTGGACCAA 2106
Qy 953 HisAlaValThrGlyIleIleValGlyValCysLeuGlyLeuLeuCysLeuAlaCys 972
Db 2107 AAATCAATGACTGGCATGCTGTAGGTGTGGCATAGCTTGACCTTGACATCTCATCTGT 2166
Qy 973 MetCysAlaGlyLeuArgGlnSerSerHisArgGluAla----- 985
Db 2167 GTTCTCATCTTGATATACCGAAGTAAAGCCAGGAATCATCTGCTTCCAAGCGGCACAG 2226
Qy 986 -----LeuProGlyLeuSerSerSerGlyThrProGlyAsn 997
Db 2227 AATGGAACTCAACAGATTACCTCGTACAGTGCCTCTTAGCTAGTGGAAAT 2277
```

RESULT 9

```
US-09-877-730-19
; Sequence 19, Application US/09877730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 2724
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-19
```

```
Alignment Scores:
Pred. No.: 1.36e-78 Length: 2724
Score: 1329.00 Matches: 312
Percent Similarity: 46.29% Conservative: 131
```

Best Local Similarity: 32.60%				Mismatches: 276			
Query Match: 20.12%				Indels: 238			
DB: 4				Gaps: 11			
US-09-754-997A-2 (1-1252) x US-09-877-730-19 (1-2724)							
Qy	79	GluHisGluAenLeuHisLeuLeuProAenGlySerLeuTrpLeuSerSerProLeuGlu	98	Qy	417	AlaProLeuAlaValValValArgGluGlyLeuProSerAlaProThrArgValThrAla	436
Db	7	GAATAAATACCGATCGAGGTCTTTCTTAACGGCTCTTTATACATCAGT-----GAG	57	Db	988	GCCAGACTGACTGTAGTGTGTGAGAGACAGACAGAGCTCTCCCTATATATGATGCT	1047
Qy	99	GlnGluAspSerAspAspGluGluAlaLeuArgIleTrpLysValThrGluGlySerTyr	118	Qy	437	ThrProLeuSerSerSerSerValLeuValAlaTrpGluArgProGluLeuHisSerGlu	456
Db	58	GTGAAGCAGCGGAGGAGCAGTCC-----GATGAAGGATTATTTAT	99	Db	1048	GAACACATGTCAAGCTCAGCATCTTTTAGCCCTGGGAGAGGCCACTTTATAAATTCAGAC	1107
Qy	119	SerCysLeuAlaHisSerProLeuGlyValValAlaSerGlnValAlaValValLysLeu	138	Qy	457	GlnIleIleGlyPheSerLeuHisTyrGlnLysAlaArgGlyValAspAenValGluTyr	476
Db	100	CAGTGTGGCAATGACAAATATAGGCCATCTTAGTCAAAAGCTCATCTTCGCCCTTA	159	Db	1108	AAATCATTCCTTATCTGTACACTACATGAAGACAGAGGTTTAAATATGAAGAGTAT	1167
Qy	139	AlaThrLeuGluAenPheSerLeuHisProGluSerGlnIleValGluGluAenGlyThr	158	Qy	477	GlnPheAlaValAenAenAspThrThrGluLeuGlnValArgAspLeuGluProAenThr	496
Db	160	TCAACTATTTCTGCATTTGAAGTCCAGCCCAATTTCCACTGAGGTCACGAAGGTGGAGTT	219	Db	1168	CAAGTAGTCATCGGAATATGACACACTCATATATATATGATGATGATGAGCCCTGCCAGC	1227
Qy	159	AlaArgPheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAsp	178	Qy	497	AspTyrGluPheTyrValValAlaTyrSerGlnLeuGlyAlaSerArgThrSerSerPro	516
Db	220	GCTCGATTGTCATGCAAGATTTCATCCACCCTCTCTGCAGTCATAACATGGAGTTCAT	279	Db	1228	AAATTATCTTCTACATTTAGCATATATGCCAATGGGAGCCAGCAGATGCTCTGACCAT	1287
Qy	179	GlnValThrValPro---GluGluProArgLeuIleThrLeuProLysTrpLeuGln	197	Qy	517	AlaLeuValHisThrLeuAspAspValProSerAlaAlaProGlnLeuThrLeuSerSer	536
Db	280	CGGCACTCTTACCTATGACTATGACAGAGTAATCTGCCCTTACCACAGGAGTATTCGAG	339	Db	1288	GTGACACAGAAATATCTAGAGGATGACCCAGCA-----	1320
Qy	198	IleLeuAspValGlnAenSerAspAlaGlySerTyrArgCysValAlaThrAsnSerAla	217	Qy	537	ProAenProSerAspIleArgValAlaTrpLeuProLeuProSerSerLeuSerAsnGly	556
Db	340	ATCTATGATGTCAGCCAAAGGATCTCGAAATATTCGTTGATGTGTCGCCACTGTAGCC	399	Db	1320	-----	1320
Qy	218	ArgGlnArgPheSerGlnGluAserLeuThrValAlaLeuArgGlySerLeuGluAla	237	Qy	557	GlnValLeuLysTyrLysIleGluTyrGlyLeuGlyLysGluAspGlnValPheSerThr	576
Db	400	CACCGAGTAAAGTATGAGGCTCGCTAACTGTGATT-----CCAGCTAAGGAGTCA	453	Db	1320	-----	1320
Qy	238	ThrArgGlyGlnAspValValIleValAlaAlaProGluAenThrThrValValSerGly	257	Qy	577	GluValProGlyAenGluThrGlnLeuThrLeuAenSerLeuGlnProAenLysValTyr	596
Db	454	AAATCTCTCCACACACCACCAATATATAGCAGGTCCACAGAACATAACACATCTCTTCAT	513	Db	1320	-----	1320
Qy	258	GlnAenValValMetGluCysValAlaSerAlaAspProThrProPheValSerTrpVal	277	Qy	597	ArgValArgIleSerAlaGlyThrGlyAlaGlyTyrGlyValProSerGlnTrpMetGln	616
Db	514	CAGACTGTAGTTTGGATGATGATGACACAGGAATCCCAACCAATCAITTCITGGAGC	573	Db	1320	-----	1320
Qy	278	ArgGlnAspGlyLysProIleSerThr---AspValIleValLeuGlyArgThrAsnLeu	296	Qy	617	HisArgThrProGlyValHisAsnGlnSerHisValProPheAlaProAlaGluLeuLys	636
Db	574	CGCCTTGATCAAAATCCATTCATGCTTTTAATATCTCGGTACTTGGAAATGGTAATCTC	633	Db	1320	-----	1320
Qy	297	LeuIleAlaSerAlaGlnProArgHisSerGlyValTyrValCysArgAlaAenLysPro	316	Qy	637	ValArgAlaLysMetGluSerLeuValValSerTrpGlnProProHisProThrGln	656
Db	634	ATGATATCTGATGTCAGCTACACATGCTGGAGTATATGTTGTCGGGCCACTACCCCT	693	Db	1320	-----	1320
Qy	317	LeuThrArgAspPheAlaThrAlaAlaGluLeuArgValLeuAlaAlaProAlaIle	336	Qy	657	IleSerGlyTyrLysLeuTyrTrpGlyGluValGlyThrGluGluGluAlaAspGlyAsp	676
Db	694	GGCACGCACTTTACAGTTGCTATGGCACTTAACTGTATTAGTCTCTCTTCATTT	753	Db	1320	-----	1320
Qy	337	SerGlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAlaArgPheValCysArg	356	Qy	677	ArgProProGlyGlyArgGlyAspGlnAlaTrpAspValGlyProValArgLeuLysLys	696
Db	754	GTTGAATGGCCAGAAAGTTTAAACAGGCCCTCGAGCTGGCTCGATTGTGTGTGTCAG	813	Db	1320	-----	1320
Qy	357	AlaSerGlyGluProArgProAlaLeuHisTrpLeuHisAspGlyIleProLeuArgPro	376	Qy	697	LysValLysGlnTyrGluLeuThrGlnLeuValProGlyArgProTyrGluValLysLeu	716
Db	814	GCAGAAGGAATCCCTCTCCCAAGATGTCATGTTGAAATATGGAAGGAGACATTCG	873	Db	1321	-----	1321
Qy	377	AsnGlyArgValLysValGlnGlyGlySerLeuValIleThrGlnIleGlyLeu	396	Qy	717	ValAlaPheAsnLysHisGluAspGlyTyrAlaAlaValTrpLysGlyLysThrGluLys	736
Db	874	AATGGTAGAATTAATAATGTAC-----AACAGTAAATTTGGTAATTAACAGATTATTCT	927	Db	1342	CTGGCTTACACAAACATAGACGATGGCTATCAGCA-----GATCAGACT	1386
Qy	397	GlnAspAlaGlyTyrTyrGlnCysValAlaGluAenSerAlaGlyThrAlaCysAlaAla	416	Qy	737	AlaProThrProAspLeuProIleGlnArgGly-----ProProLeuProPro	752
Db	928	GAAGATGATGCTATTTATCAGTGTGCTGAGGAAATAGCCAGGATCTATTTTATCTAGA	987	Db	1387	GTGAGCACTCAGAGTCGGTGTCTGTCGTGATCGCATGGTCTCTCTCCACCAACCC	1446

829 ATTTCTGTGAGTGGCAGCAATGTAGAGACACAGCTGCTATTTCAGGGCTACAGCTG 888
Qy
663 TyrTrpGlyGluValGlyThrGluGluAlaAspGlyAspArgProGlyGlyArg 682
Db
889 TACTACAGGAAGAGGAGCAGAGAAAT-----918
Qy
663 GlyAspGlnAlaTrpAspValGlyProValArgLeuLysLysValLysGlnTrpGlu 702
Db
919 -----GGGCCCATTTTCTTGGATACCAAGGACCTACTCTATATCT 957
Qy
703 LeuThrGlnLeuValProGlyArgProTyrGluValLysLeuValAlaPheAsnLysHis 722
Db
958 CTCAGTGGCTTAGACCCAGAGAAATATCATGTGAGACTCTCGCTTACACACATA 1017
Qy
723 GluAspGlyTyrAlaValTrpLysGlyLysThrGluLysAlaProThrProAspLeu 742
Db
1018 GACGATGGCTATCAGGCA-----GATCAGACTGTCCAGCACTCCAGGATGC 1062
Qy
743 ProIleGlnArgGly-----ProProLeuProProAlaHisValHisAlaGlu 758
Db
1063 GTGCTGTGGTGGATCGCATGCTCCTCCACCAACCCACCATCTCTATCGAAG 1122
Qy
759 SerAsnSerThrSerIleTrpLeuArgTrpLysLysProAspPheThrThrValLys 778
Db
1123 GCTAACACTCATCTTCCATCTTCTGCTGAGGAGGCTGCAATTCACGCTGCACA 1182
Qy
779 IleValAsnTyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThr 798
Db
1183 ATCATTAATACTACACCATCGCTGTAATCTGTGTGGCTGCAGATGCTTCTTTGGTCTG 1242
Qy
799 TyrTyrThrSerSerGlyGluAspIleLeuLysLysLeuLysProPheThrThrLysTyr 818
Db
1243 TACCTTCAACACATCAAGAACTCACATGTGTGGTTCAGAGTCTAGAACCAACCAATAC 1302
Qy
819 GluPheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGlu 838
Db
1303 GAATTTGGCTGCGATTATCATGTGATCAGCTTTCCAGTCTTGGAGCCCTGTAGTCTAC 1362
Qy
839 ArgSerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeu 858
Db
1363 CATCTACTCTTCCAGAACACACAGCCAGCCACCCAGTGGAGTAAAGTGACATTAATA 1422
Qy
859 ThrProSerThrValArgLeuHisTrpCysProProThrGluProAsnGlyGluLeuVal 878
Db
1423 GAGGATGACATCGCCCTGCTTTCTTGGAAACCCCTGTATGGCCAGAACAGTGTGACC 1482
Qy
879 GluTyrLeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeu 898
Db
1483 CGCTATATCTTATATGCACTAGGAGGCTGATTTGCGAGGAGTGGCAGGTCTTAA 1542
Qy
899 ThrThrGluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyr 918
Db
1543 CACCGTGAAGGGCAATAACCATGCTTGTCTAGAAAACCTTGGTAGCAGGAAATGTGTAC 1602
Qy
919 PhePheLysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuAsp 938
Db
1603 ATTTGCAAGATCTGCTCAATCAGGTGGGAGGAGGACCCCTTTCAAAATTCGTGGAG 1662
Qy
939 ValIleThrLeuGlnGluThrPheSerAspSer-----949
Db
1663 CTGGCAGTACTTCCAAAGGAACCTCTGAATCAAAATCAGAGCCCAAGCGTTTAGATTCT 1722
Qy
950 -----LeuAspValHisAlaValThrGlyIle 958
Db
1723 GCTGATGCCAAGTTTATTCAGGATATTACCATCTGGACCAAAATCAATGACTGCCATT 1782
Qy
959 IleValGlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArg 978
Db
1783 GCTGTAGGTGTGGCATAGCTTGCCTTGCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 1842
Qy
979 GlnSerSerHisArgGluAla-----Leu 986
Db
1843 CGAAGTAAAGCAGGAAATCATCTGCTTCCAGACGGCACAGAAATGGAACCTCAACAGTTA 1902

Qy 987 ProGlyLeuSerSerSerGlyThrProGlyAsn 997
Db 1903 CCTGTACCAAGTGCCTCTTGTAGCTAGTGAAT 1935

RESULT 11

US-09-877-730-27
; Sequence 27, Application US/09877730
; Patent No. 6465632
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6465632el Human Phosphatases and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0189-USA
; CURRENT APPLICATION NUMBER: US/09/877,730
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,607
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 2382
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-877-730-27

Alignment Scores:
Pred. No.: 1.85e-66 Length: 2382
Score: 1144.00 Matches: 252
Percent Similarity: 54.69% Conservative: 115
Best Local Similarity: 37.56% Mismatches: 240
Query Match: 17.32% Indels: 64
Gaps: 10

US-09-754-997A-2 (1-1252) x US-09-877-730-27 (1-2382)

Qy 365 LeuHisTrpLeuHisAspGlyIleProLeuArgProAsnGlyArgValLysValGlnGly 384
Db 1 ATGTCATGTTTGAATAATGGAAGGAGATACATTCGAATGGTAGAATTTAAATGTAC--- 57
Qy 385 GlyGlyGlySerLeuValIleThrGlnIleGlyLeuGlnAspAlaGlyTyrGlnCys 404
Db 58 ---AACAGTAATTTGGTAATTAACCAATATTCTCTGAAGATGATGCTATTATTCAGTGC 114
Qy 405 ValAlaGluAsnSerAlaGlyThrAlaCysAlaAlaProLeuAlaValValArg 424
Db 115 ATGGCTGAGATAGCCCAAGGATCTATTATCTAGAGCCAGACTGACTGTAGTGTCA 174
Qy 425 GluGlyLeuProSerAlaProThrArgValThrAlaThrProLeuSerSerSerVal 444
Db 175 GAACAGACAGCCAGTGTCTCCCTATATATGTATACATGCTGAACCATGTCTCAAGCTCAGCCATT 234
Qy 445 LeuValAlaTrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHis 464
Db 235 CTTTATGCTGGAGAGGCCACTTTATTAATTCAGCAAGTCAATTCCTATTCGTACAC 294
Qy 465 TyrGlnLysAlaArgGlyValAspAsnValGluTyrGlnPheAlaValAsnAspThr 484
Db 295 TACATGAAGCAGAGGTTTAAATTAATGAAGATATCAAGTAGTCACTCGGAATGACACA 354
Qy 485 ThrGluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPheTyrValValAla 504
Db 355 ACTCATTAATATTATGATGACTTAGAGCCTCCAGCAATTAATCTTCTACATGTAGCA 414
Qy 505 TyrSerGlnLeuGlyAlaSerArgThrSerProAlaLeuValHisThrLeuAsp 524
Db 415 TATATGCCAATGGGAGCCAGCCAGTGTCTGACCATGTGTGACACAGAACTACTTAGAGAT 474

Qy 525 ValProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAspIleArgVal 544
Db 475 GTTCCCTGAGACCTCTGAAATAGTTGACAGTGAAGTCCCACTGATATCTCATC 534
Qy 545 AlaTrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLysTyrIleGlu 564
Db 535 TCCTGGCTGCCAATCCAGCAATATCGCGGGGCAAGTGGTGTATGCTTGTCT 594
Qy 565 TyrGlyLeuGlyLysGluAspGlnValPheSerThrGluValProGlyAsnGluThrGln 584
Db 595 TTCGGCTAGTACTGAGAAATCAATCAAGTCTCGAGTCTCCGGGACCCAGCATGAG 554
Qy 585 LeuThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgIleSerAlaGlyThr 604
Db 655 TACCTTTGGAGGCTGGAACCTGACAGTGTCTACTGCTCGGATTAAGTGTGCCACC 714
Qy 605 GlyAlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrProGlyValHisAsn 624
Db 715 AGAGTGGGGCTGGGAGAGTCAATGATGACTTCAATAGGACGCGCCAAAGCT---ACA 771
Qy 625 GlnSerHisValProPheAlaProAlaGluLeuLysValArg---AlaLysMetGluSer 643
Db 772 AGCGTGAAGCCCTTAAGTCTCCA---GAGTTGCAATTTGGAGCTCTGAACTGTACACC 828
Qy 644 LeuValValSerTrpGlnProProHisProThr---GlnIleSerGlyTyrLysLeu 662
Db 829 ATTCTGTGAGTGGGAGCAGAGATGTAGAGGACACAGCTCTATTGCGGCTACAGCTG 888
Qy 663 TyrTrpGlyGluValGlyThrGluGluGluAlaAspGlyAspArgProGlyGlyArg 682
Db 889 TACTACAAGGAAGGAGGCGCAGAGGAAT-----GATCAGACTGTGCACATCCAGGATGC 1062
Qy 683 GlyAspGlnAlaTrpAspValGlyProValArgLeuLysLysValLysGlnTrpGlu 702
Db 919 -----GGGCCCATTTCTTGATACCAAGGACCTACTCTATCTACT 957
Qy 703 LeuThrGlnLeuValProGlyArgProTyrGluValLysLeuValAlaPheAsnLysHis 722
Db 958 CTGAGTGGCTAGACCCAGAGAAATATCATGTGAGACTCTGCTTACACACATA 1017
Qy 723 GluAspGlyTyrAlaAlaValTrpLysGlyLysThrGluLysAlaProThrProAspLeu 742
Db 1018 GACGATGGCTATCAGGCA-----GATCAGACTGTGCACATCCAGGATGC 1062
Qy 743 ProIleGlnArgGly-----ProLeuProProAlaHisValHisAlaGlu 758
Db 1063 GTGTCTCTGTGATCGATCGTCCCTCCACCCACCCACCATCTCTATGCGAAG 1122
Qy 759 SerAsnSerThrSerIleTrpLeuArgTrpLysLysProAspPheThrThrValLys 778
Db 1123 GCTAACACCTCACTTCATCTCTCGACTGAGGAGGCTGCATTCACCGCTGCACAA 1182
Qy 779 IleValAsnTyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThr 798
Db 1183 ATCATTAACATACACCATCCGCTGTAATCTGTTGGCTGCGAGATGCTTCTTGTGTTCTG 1242
Qy 799 TyrTrpThrSerSerGlyGluAspIleLeuLysGlyLysLeuLysProPheThrLysTyr 818
Db 1243 TACCTTCAACATCAGAAATCATCATGTTGGTTCAGGTCTAGAACCAACCAACATAC 1302
Qy 819 GluPheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGlu 838
Db 1303 GAATTCGGCTTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1362
Qy 839 ArgSerThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeu 858
Db 1363 CATTTCTACTCTTCCAGAGGACACAGAGGCGCCAGTGGAGTAAAGTGACATTAATA 1422
Qy 859 ThrProSerThrValArgLeuHisTrpCysProThrGluProAsnGlyGluLeuVal 878
Db 1423 GAGGATGACATGCTCCCTGTTCTTGGAAACCCCTGATGCCCAAGAAACAGTTGTGACC 1482
Qy 879 GluTyrLeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeu 898

Db 1483 CGCTATACCTATCTTATATGCTATAGGAAGCGCTGGATTCGAGGAGAGTGGCAGGTCTTA 1542
Qy 899 ThrThrGluGlyAsnIlePheSerAlaGluValHisGlyLeuLysSerAspThrArgTyr 918
Db 1543 CACCTGGAAGGGGCAATACCAATGCTTTGCTAGAAAACCTTGGTAGCAGGAAATGTGTAC 1602
Qy 919 PhePheLysMetGlyValArgThrGluValGlyProGlyProPheSerArgLeuGlnAsp 938
Db 1603 ATTGTCAAGATATCTGCAATCAATGAGTGGGAGAGGACCCCTTTTCAAATTTCTGTGGAG 1662
Qy 939 ValIleThrLeuGlnGluThrPheSerAspSer----- 949
Db 1663 CTGGCAGTACTTCCAAAGGAAACCTCTGAATCAATCAAGGCGCCCAAGCGTTTAGATTCT 1722
Qy 950 -----LeuAspValHisAlaValThrGlyIle 958
Db 1723 GCTGATGCCAAAGTTTATTTCAGGATATTACCATCTGGACCAAAAATCAATGACTGGCATT 1782
Qy 959 IleValGlyValCysLeuGlyLeuLeuCysLeuLeuAlaCysMetCysAlaGlyLeuArg 978
Db 1783 GCTGTAGGTGTGGCATAGCCTTGACCTGCACTCTCATCTGTGTTCTCATCTTGTATATAC 1842
Qy 979 GlnSerSerHisArgGluAla-----Leu 986
Db 1843 CGAAGTAAAGCCAGGAAATCATCTCTCCAAGACGCGCAGACAGATGGAACCTCAACAGTTA 1902
Qy 987 ProGlyLeuSerSerSerGlyThrProGlyAsn 997
Db 1903 CCTCGTACCAGTGCCTCTCTAGTAGTGGAAAT 1935

RESULT 12
US-09-041-886-24
; Sequence 24, Application US/09041886
; Patent No. 6235872
; GENERAL INFORMATION:
; APPLICANT: Bredesen, Dale E.
; APPLICANT: Rabizadeh, Sharroz
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence
; TITLE OF INVENTION: Polypeptides and Methods of Use
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/041,886
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2626
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4608 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS

LOCATION: 1...4342
US-09-041-886-24

Alignment Scores:

Pred. No.: 8,57e-53
Score: 943.00
Percent Similarity: 36.29%
Best Local Similarity: 24.54%
Query Match: 14.28%
DB: 3

Length: 4608
Matches: 357
Conservative: 171
Mismatch: 547
Indels: 381
Gaps: 47

US-09-754-997A-2 (1-1252) x US-09-041-886-24 (1-4608)

```
QY 51 valValLeuAspCysThrLeuGlyAlaThrAlaAlaGlyProThrArgValThrTrp 70
DB 169 GTCTCTCGACTGCTCC---GGGAGTCCGACCGAGGAGTTCAGTG---ATCAAGTGG 222
QY 71 SerLysAspGlyAspThrValLeuGluHisGluAsnLeuHisLeu-----85
DB 223 AAGAAAGATGGC-----ATTCACTGGCCTTGGGAATGGAT 258
QY 86 -----LeuProAsnGlySerLeuTrpLeuSerSerProLeuGluGlnGlu 100
DB 259 GAAAGGAAGCAGCAACTTTCAATGGGTCTCTGTATACAAACATCTTCATTC---315
QY 101 AspSerAspAspGluGluAlaLeuArgIleTrpLysValThrGluGlySerTrpCys 120
DB 316 -----AGACACCAAGCCAGATGGGAGCTTTACCAATGT 351
QY 121 LeuAlaHis---SerProLeuGlyValValAlaSerGlnValAlaValLysLeuAla 139
DB 352 GAGGCATCTTTAGGAGATCTGGCTCAATATTATTAGTGGACAGCAAAAGTTGAGTAGCA 411
QY 140 ThrLeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThrAla 159
DB 412 GGACCACTGAGGTCTTTTACAGACAGATCTGTACAGCCTTCATGGAGACAGATG 471
QY 160 ArgPheGluCysHisThrLysGlyLeuProAlaProIleIleThrTrpGluLysAspGln 179
DB 472 CTACTCAAGTGTGAAGTCAATGGGGAGCCCATGCCAACATCCACTGGCAGGAACCAA 531
QY 180 -----ValThrValProGluGluProArgLeuIleThrLeuProLysTrpLeuLeu 196
DB 532 CAAGACTGACTCCCAATCCAGGTGACTCCGAGTGGTGGTCTTGGCCCTCTGGAGCATTTG 591
QY 197 GlnIleLeuAspValGlnAspSerAlaGlySerTrpArgCysValAlaThrAsnSer 216
DB 592 CAGATCAGCGGACTCCACCGGGGGACATTGGNAATTACGATGCTCAGCTCGAATCCCA 651
QY 217 AlaArgGlnArgPheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGlu 236
DB 652 GCCAGCTCAAGACAGCAAGATGAAGCAGAGATTCAGAAATTTATCAGATCCAGGACTGCAT 711
QY 237 AlaThrArgGlyGlnAspValValIleValAlaProGluAsnThrThrValValSer 256
DB 712 -----AGACAGCTGTATTTTCTGCAAGACCATCCCAATGTAGTAGCCATTGAA 759
QY 257 GlyGlnAsnValValMetGluCysValAlaSerAlaAspProThrProPheValSerTrp 276
DB 760 GGAAGAGATGCTGCTGGATGTTGTGTTCTTGGCTATCCTCCCAAGATTTTACCTGG 819
QY 277 ValArgGlnAspGlyLysProIle-----SerThrAspValIleValLeuGly 292
DB 820 TTACGA-----GGCGAGGAGTCATCCAACTCAGGTCTAAAAAGTATTCTTTATTTGGGT 873
QY 293 ArgThrAsnLeuLeuIleAlaSerAlaGlnProArgHisSerGlyValTrpValCysArg 312
DB 874 GGAAGCAACTTGTCTATCTCAATGTGACAGATGATGACAGTGGAAATGTATATCTGTGT 933
QY 313 AlaAsnLysProLeuThrArgAspPheAlaThrAlaAlaGluLeuArgValLeuAla 332
DB 934 GTCACA-----TATAAAATGAGATATTAGTGGCTCTGCAGAGCTCAGTCTTTGGTT 987
```

```
QY 333 AlaProAlaIleSerGlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAlaArg 352
DB 988 CGCCCATGGTTTTTAATCATCTTCCAACTGTATGCTTATGAAGCATGGATATTGAG 1047
QY 353 PheValCysArgAlaSerGlyGluProArgProAlaLeuHisTrpLeuHisAspGlyIle 372
DB 1048 TTGTGAATGTACAGTCTCTGGAAAGCTGTGCCACTGTGAATGTGAAGTGAAGTGAAGT 1107
QY 373 ProLeuArgProAsnGlyArgValLysValGlnGlyGlyGlySerLeuValIleThr 392
DB 1108 GTGGTCATTCCTAGTGAATTTTTCAGATAGTG---GGAGGAAGCAACTTACGGATCTT 1164
QY 393 GlnIleGlyLeuGlnAspAlaGlyTrpTrpGlnCysValAlaGluAsnSerAlaGlyThr 412
DB 1165 GGGGTGGTGAAGTCAGATCAAGAGCTTTTATCAATGTGTGGCTGGAATATGAGGCTGAAAT 1224
QY 413 AlaCysAlaAlaAlaProLeu-----AlaValValValArgGluGlyLeu 427
DB 1225 GCCCAGACCGAGTGCACAGCTCATTTGCTCCCTAAGCTGCAATCCCAAGCTCCAGTGTCTC 1284
QY 428 ProSerAlaProThrArgValThrAlaThrProLeuSerSerSerValLeuValAla 447
DB 1285 CTTTGGCTCCCGAGAGATGTGGTCCCTGTCTTGGTTTCCAGCCGATTTGTCCGCTCAGC 1344
QY 448 TrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTrpGlnLys 467
DB 1345 TGGCGCCCACTCTGCAAGACGGAAGCAATTTCAAACTTTTCAAGCTCTTTTTC-----1398
QY 468 AlaAlaGlyValAlaAspAsnValGluTrpGlnPheAlaValAsnAsn-----AspThr 484
DB 1399 TCCAGAGAGGTGACACACAGGGAACGA-----GCATTGAATACACACAGCGCTGGGTCC 1452
QY 485 ThrGluLeuGlnValArgAspLeuGluProAsnThrAspTrpGluPheTrpValValAla 504
DB 1453 CTTTCACTCAGTCTGGAAACCTGAAGCCAGAGCCATGTACACCTTTTCAAGTGTGGCT 1512
QY 505 TyrSerGlnLeuGlyAlaSerArgThrSerPro-----516
DB 1513 TACAATGAATGGGACCGGGAGAGAGTCTTCAACCCATCAAGGTGGCCACACAGCGCTGAG 1572
QY 516 -----516
DB 1573 TTGCAAGTTTCCAGGGCCAGTAGAATAAAGCTGTATCTACCTCACCTACCTCAATT 1632
QY 516 -----516
DB 1633 CTTATTACTTGGAAACCCCTGCTTATGCAACCGTCCAGTCCAGGTTTACAGATTGTTTC 1692
QY 516 -----516
DB 1693 TGCACCTGAGGTGTCCACAGAAAAAGAACAGATATAGAGTTGTATGAGTACTTCTTATAAA 1752
QY 516 -----516
DB 1753 CTGGAAGGCTGAAAAAATTCACCAATATAGTCTTCCGATTTCTTAGCTTATAATCGCTAT 1812
QY 517 -----AlaLeuValHisThrLeuAspAspValProSerAla 528
DB 1813 GGTTCGGGGTCTCTACTGATGATATAACAGTGGTTTACACTTTCTGAGTGCAGAGTGC 1872
QY 529 AlaProGln---LeuThrLeuSerSerProAsnProSerAspIleArgValAlaTrpLeu 547
DB 1873 CGGCTCTCAGAACGTCTCCCTGGAGTGGTCAATTCAGAGAGTATCAAGTGTAGTGGCTG 1932
QY 548 ProLeuProSerSerLeuSerAsnGlyGlnValLeuLysTrpLysIleGluTrpGlyLeu 567
DB 1933 CTTCTCTCCATCAGGAACCAAAATGGATTTTATTACCGGCTATAAAATTCGACACAGAAAG 1992
QY 568 GlyLysGluAspGlnValPheSerThrGluValProGlyAsnGluThrGlnLeuThrLeu 587
DB 1993 ACGACCCGCGGGGTGAGATGGAAACACTGGAGCCAAACCAACCTCTGGTACCTA---TTC 2049
QY 588 AsnSerLeuGlnProAsnLysValTrpArgValArgIleSerAlaGlyThrGlyAlaGly 607
```


Qy 1215 oSerThrAlaIleGluAla-----1222
 Db 3996 CACAGCTGTGTCGACCAACTCACCCACTCCGAGCTTTGCTAATCTTTGCTACCTCC 4055
 Qy 1223 -ProGlySerCysGlnProLysAlaLeuCysProLeuThr-----Va 1237
 Db 4056 ACCAATGAGTCAATAGAACCGAAGTC-----CCTTACACACACTTTTGTCTCAGCC 4109
 Qy 1237 lSerProSerLeuProArgAlaProValSerSerAlaGlnVal 1251
 Db 4110 AGGCCCACTCTTCTAAGACCCATGTGAAACAGACCTCCCTT 4152

RESULT 13

PCT-US94-05277-1
 ; Sequence 1, Application PC/TUS9405277
 ; GENERAL INFORMATION:
 ; APPLICANT: Bruskin, Arthur
 ; APPLICANT: Jarosz, David E.
 ; APPLICANT: Johnson, Karen
 ; APPLICANT: Kinzler, Kenneth W.
 ; APPLICANT: Vogelstein, Bert
 ; APPLICANT: Zabrecky, James R.
 ; TITLE OF INVENTION: Antibodies Specific for DCC Gene Product
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Banner, Birch, McKie & Beckett
 ; STREET: 1001 G Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20001

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US94/05277
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kagan, Sarah A.
 ; REGISTRATION NUMBER: 32,141
 ; REFERENCE/DOCKET NUMBER: 01107.42709
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202.508.9100
 ; TELEFAX: 202.508.9299
 ; TELEX: 197430 BBMB UT
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4608 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORGANISM: Homo sapiens
 ; POSITION IN GENOME:
 ; CHROMOSOME/SEGMENT: 18q21
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..4342
 ; PCT-US94-05277-1

Alignment Scores:
 Pred. No.: 8.57e-53
 Score: 943.00
 Percent Similarity: 36.29%
 Best Local Similarity: 24.54%
 Query Match: 14.28%
 Length: 4608
 Matches: 357
 Conservative: 171
 Mismatches: 547
 Indels: 381

DB: 5 Gaps: 47
 US-09-754-997A-2 (1-1252) x PCT-US94-05277-1 (1-4608)
 Qy 51 ValValLeuAspCysThrLeuGlyAlaThrAlaAlaGlyProProThrArgValThrTyr 70
 Db 169 GTCTCTCTCAGCTGCTCC---GCGAGTCCGACGAGAGTTCAGTG---ATCAAGTGG 222
 Qy 71 SerLysAspGlyAspThrValLeuGluHisGluAsnLeuHisLeu-----85
 Db 223 AAGAAAGATGCG---ATTCATCTGGCTTGGGAATGGAT 258
 Qy 86 -----LeuProAsnGlySerLeuTyrLeuSerSerProLeuGluGlnGlu 100
 Db 259 GAAAGGAAGCAGCAACTTTCAATGGTCTCTGCTGATACAAACACATCTTCATTCC---315
 Qy 101 AspSerAspAspGluGluAlaLeuArgIleTyrLysValThrGluGlySerTyrSerCys 120
 Db 316 -----AGACACCAACAGCCAGATGAGGACCTTTACCAATGT 351
 Qy 121 LeuAlaHis---SerProLeuGlyValValAlaSerGlnValAlaValValLeuAla 139
 Db 352 GAGCATCTTTAGGAGATTCGCTCAATTATTAGTCGGACAGCAAAAGTTGCAGTAGCA 411
 Qy 140 ThrLeuGluAspPheSerLeuHisProGluSerGlnIleValGluGluAsnGlyThrAla 159
 Db 412 GGACCACTGAGGTTCTCTTCCACAGACAGAAATCTGTCCAGCTTCTATGGGAGACACAGTG 471
 Qy 160 ArgPheGluCysHisThrLysGlyLeuProAlaProIleIleThrTyrGluLysAspGln 179
 Db 472 CTACTCAAGTGTGAAGTCATTGGGGAGCCCATCCCAACATCCATGGCAGAGAACCAA 531
 Qy 180 -----ValThrValProGluProArgLeuIleThrLeuProLysTyrLeuLeu 196
 Db 532 CAAGACTGACTCAATCCAGTGACTCCGAGTGGTGTCTTGGCTCTCGAGAGCATGT 591
 Qy 197 GlnIleLeuAspValGlnAspSerAspAlaGlySerTyrArgCysValAlaThrAsnSer 216
 Db 592 CAGATCAGCCGACTCTCCAACGGGGACATTGGGAATTTTACCGATGCTCAGCTCGAAATCCA 651
 Qy 217 AlaArgGlnArgPheSerGlnGluAlaSerLeuThrValAlaLeuArgGlySerLeuGlu 236
 Db 652 GCCAGCTCAAGAACAGGAAATGAAGCAGAGATGAGAAATTTTATCAGATCCAGGACTGCAT 711
 Qy 237 AlaThrArgGlyGlnAspValValIleValAlaProGluAsnThrThrValValSer 256
 Db 712 -----AGACAGCTGTATTTCTCAAGACCATCCCAATGATGAGCCATTGAA 759
 Qy 257 GlyGlnAsnValValMetGluCysValAlaSerAlaAspProThrProPheValSerTyr 276
 Db 760 GGAAGAGATGCTCTCTGGAAATGTTGTGTTCTGGCTATCTCTCCACCAAGTTTACCTGG 819
 Qy 277 ValArgGlnAspGlyLysProIle-----SerThrAspValIleValLeuGly 292
 Db 820 TTACGA-----GCGCAGGAGATCATCCAACTCAGGTCTAAAAAGTATCTTTATTGGGT 873
 Qy 293 ArgThrAsnLeuLeuIleAlaSerAlaGlnProArgHisSerGlyValTyrValCysArg 312
 Db 874 GGAAGCAACTTGTCTTATCTCCAATGTGACAGATGATGACAGTGAATGTATACCTGTGT 933
 Qy 313 AlaAsnLysProLeuThrArgAspPheAlaThrAlaAlaAlaGluLeuArgValLeuAla 332
 Db 934 GTCACA-----TATAAAATGAGATATTAGTGGCTCTCGAGAGCTCACAGTCTTGTT 987
 Qy 333 AlaProAlaIleSerGlnAlaProGluAlaLeuSerArgThrArgAlaSerThrAlaArg 352
 Db 988 CCGCATGGTTTTTAAATCATCTCTCAACCTGTATGCTCTATGAAGACGATGATTTGAG 1047
 Qy 353 PheValCysArgAlaSerGlyGluProArgProAlaLeuHisTyrLeuHisAspGlyIle 372
 Db 1048 TTTGAATGTACAGTCTCTGGAAGACCTGTGCCACTGTGAATTTGGATGAAGAATGGAGAT 1107
 Qy 373 ProLeuArgProAsnGlyArgValLysValGlnGlyGlyGlySerLeuValIleThr 392

Db 1108 GTGGTCATCTAGTCATATATTTTCAGATAGTG-...GGAGGAAGCAACTTACCGTACTT 1164
Qy 393 GlnIleGlyLeuGlnAspIaGlyTyrGlnCysValAlaGluAsnSerAlaGlyThr 412
Db 1165 GGGGTGGTGAAGTCAGATCAAGGCTTTTATCAATGTGTGGCTGAAATGAGGCTGGAAAT 1224
Qy 413 AlaCysAlaAlaAlaProLeu-...AlaValValValArgGluGlyLeu 427
Db 1225 GCCCAGACAGTGCACAGCTCATGTCCCTAGACCTGCAATCCCAAGCTCCAGTGTCTC 1284
Qy 428 ProSerAlaProThrArgValThrAlaThrProLeuSerSerSerValLeuValAla 447
Db 1285 CCTTCGGCTCCCAAGATGTGGTCCCTGTCTTTCAGCGCATTTGTCCGTCTCAGC 1344
Qy 448 TrpGluArgProGluLeuHisSerGluGlnIleIleGlyPheSerLeuHisTyrGlnLys 467
Db 1345 TGGCGCCCACTGCAGAGCGAAAGGGAACATTCAAACTTTTCAGGCTCTTTTC-... 1398
Qy 468 AlaArgGlyValAspAsnValGluTyrGlnPheAlaValAsnAsn-...AspThr 484
Db 1399 TCCAGAGAGGTGACACACAGGGAACGA-...GCATTGATACACACAGCTGGGTCC 1452
Qy 485 ThrGluLeuGlnValArgAspLeuGluProAsnThrAspTyrGluPheTyrValValAla 504
Db 1453 CTTCACTCACTGTGGAAACCTGAAGCCAGAGCCATGTACACCTTTTCGAGTTGTGGCT 1512
Qy 505 TyrSerGlnLeuGlyAlaSerArgThrSerPro-... 516
Db 1513 TACAATGAAATGGGGACCGGAGAGAGATTCTCAACCCATCAAGGTGGCCACACAGCCTGAG 1572
Qy 516 -... 516
Db 1573 TTGCAAGTTCACGGGCGAGTAGAAGAACCTTGCAAGCTGTATCTACCTACCTCAAT 1632
Qy 516 -... 516
Db 1633 CTTATTACCTGGGAACCCCTGCTGTATGCAACGGTCCAGTCCCAAGGTTACAGATTGTTTC 1692
Qy 516 -... 516
Db 1693 TGCACCTGAGGTGTCCACAGGAAAGAACAGAAATATAGAGTTGATGGCATCTTTATAAA 1752
Qy 516 -... 516
Db 1753 CTGGAAAGGCTGAAATAATTCACCGAATATAGTCTTCGATTCTTAGCTTATATACCTAT 1812
Qy 517 -... 517
Db 1813 GGTCCGGGCTCTCTACTGATGATATAACAGTGTGTACACTTTCTCAGGTGCCAAGTGCC 1872
Qy 529 AlaProGln-...LeuThrLeuSerSerProAsnProSerAspIleArgValAlaThrLeu 547
Db 1873 CCGCCTCAGAACGTCTCCCTGGAAAGTGGTCAATTCAGGAAGTATCAAAAGTTAGCTGGCTG 1932
Qy 548 ProLeuProSerSerLeuSerAsnGlyGlnValLeuLysTyrLysIleGluTyrGlyLeu 567
Db 1933 CTTCTCCATCAGAAACAAATAGATTATTTACCGGCTATAAATTCAGACAGAAAG 1992
Qy 568 GlyLysGluAspGlnValPheSerThrGluValProGlyAsnGluThrGlnLeuThrLeu 587
Db 1993 ACGACCCGCGAGGTGAGATGGAAACACTGGAGCCCAACCACTCTGGTACCTA-...TTC 2049
Qy 588 AsnSerLeuGlnProAsnLysValTyrArgValArgIleSerAlaGlyThrGlyAlaGly 607
Db 2050 ACAGGACTGGAGAAAGGAAGTCAGTACAGTTTCCAGGTGTCCAGCCATGACAGTCAATGGT 2109
Qy 608 TyrGlyValProSerGlnTyrMetGlnHisArgThrPro-...GlyValHisAsnGlnSer 626
Db 2110 ACTGGACCACTTCCAACTGGTATCTATGTCAGAGACTCCAGAGAAATGATCTAGATGATCT 2169
Qy 627 HisValProPheAlaProAlaGluLeuLysValArgAlaLysMetGluSerLeuValVal 646

Db 2170 CAAGTTCTGTGATCAACCAAGCTCTCTTCATGTGAGGCCCCAGACTAACTGCATCATCATG 2229
Qy 647 SerTrpGlnProProProHisProThr-...GlnIleSerGlyTyrLysLeuTyrTrpGly 665
Db 2230 AGTTGGACTCTCTCCCTGACCCAAACATCGTGTGGCAGGTATATTTATTCGGTATGCG 2289
Qy 666 GluValGlyThrGluGluGluAlaAspGlyAspArgProProGlyGlyArgGlyAspGln 685
Db 2290 GTTGGGAGCCCTTACGCTGAGACA-... 2313
Qy 686 AlaTrpAspValGlyProValArgLeuLysLysLysValLysGlnTyrGluLeuThrGln 705
Db 2314 -...GTGCGTGTGGACAGCAAGCAGCGATATATTTCCATTGAGAGG 2355
Qy 706 LeuValProGlyArgProTyrGluValLysLeuValAlaPheAsnLysHisGluAspGly 725
Db 2356 TTAGAGTCAAGTTCCTATGTAATCTCCCTTAAAGCTTTTAAACATGTCGGAGAGAGA 2415
Qy 726 TyrAlaAlaValTrpLysGlyLysThrGluLysAlaProThrProAspLeuProIleGln 745
Db 2416 GTTCCTCTTTTGAAGTGCCACCACAGGTCTATAACCGATCCCACTGACCCAGTTGAT 2475
Qy 746 -...ArgGlyProProLeu 750
Db 2476 TATTATCTCTTGTGATGATTTCGCCACCTCGGTCCAGATCTCTCCACCCCATGCTC 2535
Qy 751 ProProAlaHisValHisAlaGluSerAsnSerSerThrSerIleTrpLeuArgTrp-... 769
Db 2536 CCACCAGTAGGTACAGCTGTGCTCTTACCCATGATGCTGTGAGGTGAGTGGGCA 2595
Qy 770 -...LysLysProAspPheThrThrValLysIleValAsnTyrThrVal 784
Db 2596 GACAACTCTGCTCCCTAAGAACCAAGACGCTGTGAGGTGGACTT-...TACACCGTC 2649
Qy 785 ArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTyr-... 799
Db 2650 CGG-...TGGAGAACCAAGCTTTTCTGCAAGTGCAGAAATACAAAGTCAGAGACACA 2700
Qy 800 -...TyrThrSerSerGlyGluAspIleLeuIleGlyLysLeuLysProPhe 815
Db 2701 ACATCTCTAAGTTACACAGCAACA-...GGCCTCAACCAAC 2739
Qy 816 ThrLysTyrGluPheAlaValGlnSerHisGlyValAspMetAspGlyProPheGlySer 835
Db 2740 ACAATGTATGAATCTCGGTATGTAACAAAAACAGAAAGTCCAGTACTTGGAGCATG 2799
Qy 836 ValValGluArgSerThrLeuProAspArgProSerThrProProSerAspLeuArgLeu 855
Db 2800 ACTGCACATGCCACCAAGCTATGAAGCAGCCGCCCTCTGCTCCCAAGGACTTTACAGTC 2859
Qy 856 SerProLeuThr-...ProSerThrValArgLeuHisTyrCysProProThr 871
Db 2860 -...ATTACTAGGAAGGAGGCTCTGCGCTCATTTGTGAGTTGGCAGGCTCCCTTGG 2913
Qy 872 GluProAsnGlyGluIleValGluTyrLeuIleLeuTyrSerAsnAsnHisThrGlnPro 891
Db 2914 GAAGCAATGGGAATTTACTCTTACATCTTATTTATACCTTGGACAAAGAACATCCCA 2973
Qy 892 GluHisGlnTrpThrLeuLeuThrGluGlyAsnIlePheSerAlaGluValHisGly 911
Db 2974 ATTGATGACTGTGATTTGAAACAATCAGTGTGTAGGCTTACTCATCAATCATGAT 3033
Qy 912 LeuGluSerAspThrArgTyrPhePheLysMetGlyAlaArgThrGluValGlyProGly 931
Db 3034 CTCACCTTGATCTATGTATTACTTTTGAATTCAAGCAGCAAAATTCAAAGAGGTGGG 3093
Qy 932 ProPheSer-... 934
Db 3094 CCATCTCTGATCCCATCTCTTTCAGACTCTGAAAGTGGAAACACCTGCACAAATGGCT 3153
Qy 935 -...-ArgLeuGlnAsp 938
Db 3154 AATGACCAAGGTGTCATGGAGATGGAGGTATTATGGCCAGTTTGATTAATTTGATGAT 3213


```

QY 505 TyrSerGlnLeuGlyAlaSerArgThrSerSerProAlaLeuValHisThrLeuAsp 524
Db 415 TATATGCCAATGGAGCCAGCATGCTGACCATGTGCACACAGATACCTAGAGCAT 474
QY 525 ValProSerAlaAlaProGlnLeuThrLeuSerSerProAsnProSerAspIleArgVal 544
Db 475 GACCCACAGA----- 483
QY 545 AlaTrpLeuProLeuProSerSerLeuSerAsnGlyGlnValLeuLysTyrLysIleGlu 564
Db 483 ----- 483
QY 565 TyrGlyLeuGlyLysGluAspGlnValPheSerThrGluValProGlyAsnGluThrGln 584
Db 483 ----- 483
QY 585 LeuThrLeuAsnSerLeuGlnProAsnLysValTyrArgValArgIleSerAlaGlyThr 604
Db 483 ----- 483
QY 605 GlyAlaGlyTyrGlyValProSerGlnTrpMetGlnHisArgThrProGlyValHisAsn 624
Db 483 ----- 483
QY 625 GlnSerHisValProPheAlaProAlaGluLeuLysValArgAlaLysMetGluSerLeu 644
Db 483 ----- 483
QY 645 ValValSerTrpGlnProProHisProThrGlnIleSerGlyTyrLysLeuTyrTrp 664
Db 483 ----- 483
QY 665 GlyGluValGlyThrGluGluGluAlaAspGlyAspArgProProGlyGlyArgGlyAsp 684
Db 483 ----- 483
QY 685 GlnAlaTrpAspValGlyProValArgLeuLysLysLysValLysGlnTyrGluLeuThr 704
Db 483 ----- 483
QY 705 GlnLeuValProGlyArgProTrpGluValLysLysLeuValAlaPheAsnLysHisGluAsp 724
Db 484 -----AGAAAATATCATGTGAGACTCTCGGCTTACAAACATAGACGAT 528
QY 725 GlyTyrAlaAlaValTrpLysGlyLysThrGluLysAlaProThrProAspLeuProIle 744
Db 529 GGCTATCAGGCA-----GATCAGACTGTCCAGCACTCCAGGATGCGTGCT 573
QY 745 GlnArgGly-----ProProLeuProProAlaHisValHisAlaGluSerAsn 760
Db 574 GTTCGTGATCGCATGGTCCCTCTCCACCACCCACCCTCTCTATGCGAAGGCTAAC 633
QY 761 SerSerThrSerIleTrpLeuArgTrpLysLysProAspPheThrThrValLysIleVal 780
Db 634 ACCTCATCTTCCATCTCTGCACTGGAGAGGCTGCACTTCCCGCTGCACAAATCAT 693
QY 781 AsnTyrThrValArgPheGlyProTrpGlyLeuArgAsnAlaSerLeuValThrTyr 800
Db 694 AACTACACCATCCGCTGTAATCTCTGTGGCTGCAGAAATGCTTCTTGTGTTCTGTACCTT 753
QY 801 ThrSerSerGlyGluAspIleLeuGlyGlyLeuLysProPheThrLysTyrGluPhe 820
Db 754 CAAACATCAGAACTCACATGTTGGTTCAAGTCTAGAACCCAAACCAACCAATACGAATTT 813
QY 821 AlaValGlnSerHisGlyValAspMetAspGlyProPheGlySerValValGluArgSer 840
Db 814 GCGGTTCGATTACATGTGATCAGCTTTCAGTCTCTGGAGCCCTGTAGTCTTACCATTCT 873
QY 841 ThrLeuProAspArgProSerThrProProSerAspLeuArgLeuSerProLeuThrPro 860
Db 874 ACTCTTCCAGAGACCCAGCAGGCCCCACCATGTTGGAGTAAAGAGTGCATTAATAGAGGAT 933

```

```

QY 861 SerThrValArgLeuHisTrpCysProProThrGluProAsnGlyGluIleValGluTyr 880
Db 934 GACACTGCCCTGGTTCTTGGAAACCCCTGATGGCCCAAGAAACAGTTGTGACCCGCTAT 993
QY 881 LeuIleLeuTyrSerAsnAsnHisThrGlnProGluHisGlnTrpThrLeuLeuThrThr 900
Db 994 ACTATCTTATATGATCTAGGAAGCCCTGGATTGCCAGGAGTGGCAGGTCTTACACCGT 1053
QY 901 GluGlyAsnIlePheSerAlaGluValHisGlyLeuGluSerAspThrArgTyrPhePhe 920
Db 1054 GAAGGGGCAATACCATGGCTTTGCTAGAAAACCTTGGTAGCAGGAATGTGTACATTGTC 1113
QY 921 LysMetGlyAlaArgThrGluValGlyProGlyProPheSerArgLeuGlnAspValIle 940
Db 1114 AAGATATCTGCATCCAATGAGGTGGAGAGGACCCCTTTTCAAATTCGTGGAGCTGGCA 1173
QY 941 ThrLeuGlnGluThrPheSerAspSer----- 949
Db 1174 GTACTTCCAAAGGAAACCTCTGAATCAATCAGAGGCCCAAGCGTTTAGATTCTGCTGAT 1233
QY 950 -----LeuAspValHisAlaValThrGlyIleIleVal 960
Db 1234 GCCAAAGTTTATTTCAGGATATTACCATCTGGACCACCAAAATCAATGACTGGCATTTGCTGTA 1293
QY 961 GlyValCysLeuGlyLeuLeuCysLeuAlaCysMetCysAlaGlyLeuArgGlnSer 980
Db 1294 GGTGTTGGCATAGCTTGGACCTGCATCTCATCTGTGTCTTCTCATCTTTGATATACCGAAT 1353
QY 981 SerHisArgGluAla-----LeuProGly 988
Db 1354 AAAGCCAGGAATCATCTGCTTCCAGAGGCCACAGAAATGGAACTCAACAGTTTACCTCGT 1413
QY 989 LeuSerSerSerGlyThrProGlyAsn 997
Db 1414 ACCAGTGCCTCCTTAGCTAGTGGAAAT 1440

```

Search completed: October 14, 2003, 09:29:47
Job time : 274 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2003, 15:48:26 ; Search time 10461.9 Seconds
(without alignments)
10933.318 Million cell updates/sec

Title: US-09-754-997A-3

Perfect score: 2796

Sequence: 1 ggggagtgccattgcacca.....cagactccttgatgtgcac 2796

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_em.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

42: em_htgo_mus.*

43: em_htgo_mus.*

44: em_htgo_mus.*

45: em_htgo_mus.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2796	100.0	2796	6	AX191261	AX191261 Sequence
2	2796	100.0	3756	6	AX191303	AX191303 Sequence
3	2796	100.0	6176	6	AX191259	AX191259 Sequence
4	2796	100.0	6176	10	AF176694	AF176694 Mus muscu
5	2775.2	99.3	6219	10	AB052820	AB052820 Mus muscu
6	2762.2	98.8	6222	10	AB052821	AB052821 Mus muscu
7	2595.2	92.8	6301	10	AB122535	AB122535 Mus muscu
8	2159.2	77.2	6485	9	AB052622	AB052622 Homo sapi
9	2090.8	74.8	3741	6	AX418406	AX418406 Sequence
10	1491.4	53.3	5549	9	AB046848	AB046848 Homo sapi
C 11	415	14.8	116580	10	AC112162	AC112162 Mus muscu
C 12	415	14.8	158055	2	AC110235	AC110235 Mus muscu
C 13	376.8	13.5	224227	2	AC106509	AC106509 Rattus no
C 14	331.2	11.8	176282	2	AC068507	AC068507 Homo sapi
15	331.2	11.8	198295	9	AC105129	AC105129 Homo sapi
16	316.8	11.3	2967	9	BC042054	BC042054 Homo sapi
17	316.6	11.3	2976	6	AR237558	AR237558 Sequence
18	316.6	11.3	3210	6	AR237553	AR237553 Sequence
19	316.6	11.3	3219	6	AR237561	AR237561 Sequence
20	316.6	11.3	3874	6	AR237568	AR237568 Sequence
21	316.4	11.3	2700	6	AX497857	AX497857 Sequence
22	315.4	11.3	3453	6	AR237556	AR237556 Sequence
23	302.8	10.8	3198	10	BC053057	BC053057 Mus muscu
24	301.2	10.8	3146	10	AF026465	AF026465 Mus muscu
25	288	10.3	288	6	AX191277	AX191277 Sequence
26	287.4	10.3	2729	9	AK095529	AK095529 Homo sapi
C 27	252.6	9.0	115135	10	AC112161	AC112161 Mus Muscu
28	252	9.0	252	6	AX191281	AX191281 Sequence
29	249	8.9	249	6	AX191273	AX191273 Sequence
30	249	8.9	249	6	AX191275	AX191275 Sequence
31	246	8.8	246	6	AX191279	AX191279 Sequence
32	243	8.7	243	6	AX191265	AX191265 Sequence
33	217	7.8	2481	6	AR237560	AR237560 Sequence
34	217	7.8	2715	6	AR237555	AR237555 Sequence
35	217	7.8	2724	6	AR237562	AR237562 Sequence
36	217	7.8	2958	6	AR237557	AR237557 Sequence
C 37	215.8	7.7	147760	9	AC011846	AC011846 Homo sapi
38	215.8	7.7	176282	2	AC068507	AC068507 Homo sapi
39	213.8	7.6	2139	6	AR237563	AR237563 Sequence
40	213.8	7.6	2382	6	AR237566	AR237566 Sequence
41	195	7.0	195	6	AX191271	AX191271 Sequence
42	192	6.9	192	6	AX191267	AX191267 Sequence
43	189	6.8	189	6	AX191269	AX191269 Sequence
C 44	158.8	5.7	256608	2	AC021040	AC021040 Homo sapi
45	158	5.7	158055	2	AC110235	AC110235 Mus muscu

ALIGNMENTS

RESULT 1	AX191261	Sequence 3 from Patent WO0149714.	2796 bp	DNA	linear	PAT 15-AUG-2001
LOCUS	AX191261	Sequence 3 from Patent WO0149714.	2796 bp	DNA	linear	PAT 15-AUG-2001
DEFINITION	AX191261	Sequence 3 from Patent WO0149714.	2796 bp	DNA	linear	PAT 15-AUG-2001
ACCESSION	AX191261	Sequence 3 from Patent WO0149714.	2796 bp	DNA	linear	PAT 15-AUG-2001
VERSION	AX191261.1	GI:15209512	2796 bp	DNA	linear	PAT 15-AUG-2001
KEYWORDS	AX191261.1	GI:15209512	2796 bp	DNA	linear	PAT 15-AUG-2001
SOURCE	AX191261.1	GI:15209512	2796 bp	DNA	linear	PAT 15-AUG-2001
ORGANISM	AX191261.1	GI:15209512	2796 bp	DNA	linear	PAT 15-AUG-2001
REFERENCE	AX191261.1	GI:15209512	2796 bp	DNA	linear	PAT 15-AUG-2001
AUTHORS	AX191261.1	GI:15209512	2796 bp	DNA	linear	PAT 15-AUG-2001
TITLE	AX191261.1	GI:15209512	2796 bp	DNA	linear	PAT 15-AUG-2001
JOURNAL	AX191261.1	GI:15209512	2796 bp	DNA	linear	PAT 15-AUG-2001
Salbaum, J.M.	Salbaum, J.M.	Salbaum, J.M.	Salbaum, J.M.	Salbaum, J.M.	Salbaum, J.M.	Salbaum, J.M.
Nope polypeptides, encoding nucleic acids and methods of use	Nope polypeptides, encoding nucleic acids and methods of use	Nope polypeptides, encoding nucleic acids and methods of use	Nope polypeptides, encoding nucleic acids and methods of use	Nope polypeptides, encoding nucleic acids and methods of use	Nope polypeptides, encoding nucleic acids and methods of use	Nope polypeptides, encoding nucleic acids and methods of use
Patent: WO 0149714-A 3 12-JUL-2001;	Patent: WO 0149714-A 3 12-JUL-2001;	Patent: WO 0149714-A 3 12-JUL-2001;	Patent: WO 0149714-A 3 12-JUL-2001;	Patent: WO 0149714-A 3 12-JUL-2001;	Patent: WO 0149714-A 3 12-JUL-2001;	Patent: WO 0149714-A 3 12-JUL-2001;

Pred. No. is the number of results predicted by chance to have a

601 AGCCAGAGGCGCTCGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTACAGAGGGGCGAG 660
664 AGCCAGAGGCGCTCGCTCACTGTGGCCCTCAGAGGGTCTTTGGAGGCTACAGAGGGGCGAG 723
661 GATGTGCTATTTGTGGAGGCGCCAGAGAACACACAGGTAGTGTCTGGAGCAGAGATGTAGTG 720
724 GATGTGCTATTTGTGGAGGCGCCAGAGAACACACAGGTAGTGTCTGGAGCAGAGATGTAGTG 783
721 ATGGAGTGTGGTGGCTCTGTGACCCCAACCCCTTTTGTGCTGTGGTCCGACAGATGGA 780
784 ATGGAGTGTGGTGGCTCTGTGACCCCAACCCCTTTTGTGCTGTGGTCCGACAGATGGA 843
781 AAGCCCTATCTCCACGGATGTCTATCTGTGGGCGGACCAATCTACTCATCGCCAGCGCG 840
844 AAGCCCTATCTCCACGGATGTCTATCTGTGGGCGGACCAATCTACTCATCGCCAGCGCG 903
841 CAGCCTCGGCACTCTGAGTGTATGTCTGCGAGCGCAACAGGCCCTCTACCGGTGACTTC 900
904 CAGCCTCGGCACTCTGAGTGTATGTCTGCGAGCGCAACAGGCCCTCTACCGGTGACTTC 963
901 GCACCTGGGGCTGTGAGCTCCGAGTGTCTGTGCTGCCAGCCATCTCGCAGGCAACCCGAG 960
964 GCACCTGGGGCTGTGAGCTCCGAGTGTCTGTGCTGCCAGCCATCTCGCAGGCAACCCGAG 1023
961 GCGCTCTCGCGACCGCGGCGCAGCACCGCGCTTCTGTGTGCGGGCGTCCGGGAGCCA 1020
1024 GCGCTCTCGCGACCGCGGCGCAGCACCGCGCTTCTGTGTGCGGGCGTCCGGGAGCCA 1083
1021 CGGCCCGCGCTGCACTGGCTGCAAGCGGGATCCCGTTGCGACCCCAATGGGGCGGTCAAG 1080
1084 CGGCCCGCGCTGCACTGGCTGCAAGCGGGATCCCGTTGCGACCCCAATGGGGCGGTCAAG 1143
1081 GTGCGAGGGGGTGGGGAGCTTGTGTCACTCAGATCGGCTCGCAGGACGCTGCGCTAC 1140
1144 GTGCGAGGGGGTGGGGAGCTTGTGTCACTCAGATCGGCTCGCAGGACGCTGCGCTAC 1203
1141 TACCACTGTGTAGCAGAGAAACAGCGCGGAACTGCTGTGCGGCTGCGGCCCTGGCGGTA 1200
1204 TACCACTGTGTAGCAGAGAAACAGCGCGGAACTGCTGTGCGGCTGCGGCCCTGGCGGTA 1263
1201 GTGGTGTGCGAGGGGTGCGCAGCGCCCGCACTGCGGTTCAAGCAACCGCGCTGAGCAGC 1260
1264 GTGGTGTGCGAGGGGTGCGCAGCGCCCGCACTGCGGTTCAAGCAACCGCGCTGAGCAGC 1323
1261 TCCTGTGTGCTGTGGCTGTGGAGCGGCTGAGTGTGCAAGCGAGCAATCAATGCGCTTC 1320
1324 TCCTGTGTGCTGTGGCTGTGGAGCGGCTGAGTGTGCAAGCGAGCAATCAATGCGCTTC 1383
1321 TCCTGTGTGCTGTGGCTGTGGAGCGGCTGAGTGTGCAAGCGAGCAATCAATGCGCTTC 1380
1384 TCCTGTGTGCTGTGGCTGTGGAGCGGCTGAGTGTGCAAGCGAGCAATCAATGCGCTTC 1443
1381 AATGACACACAGAGCTGAGGTGTGGGACCTGGAAACCCAAACAGGATTAAGTTCATAC 1440
1444 AATGACACACAGAGCTGAGGTGTGGGACCTGGGACCTGGAAACCCAAACAGGATTAAGTTCATAC 1503
1441 GTGGTGGCTATCTCCAGCTGGGGGCGCAGCGGAACTTCCAGCGCAGCGCTGTGATACA 1500
1504 GTGGTGGCTATCTCCAGCTGGGGGCGCAGCGGAACTTCCAGCGCAGCGCTGTGATACA 1563
1501 GTGGAGCATGTCTCCAGCGCAGCACCCAGCTTACCTGTCCAGCGCCCAACCCCTCGGAC 1560
1564 GTGGAGCATGTCTCCAGCGCAGCACCCAGCTTACCTGTCCAGCGCCCAACCCCTCGGAC 1623
1561 ATCAGGGTGGCATGTGCTGCCCTCGCCTCAGCCTGAGCAATGGACAGGTGTGAAAGTAC 1620
1624 ATCAGGGTGGCATGTGCTGCCCTCGCCTCAGCCTGAGCAATGGACAGGTGTGAAAGTAC 1683
1621 AAGATAGATACGGTTTGGGGAGGAGATCAGGTTTTCTCCACCGAGGTGCTGGAAT 1680
1684 AAGATAGATACGGTTTGGGGAGGAGATCAGGTTTTCTCCACCGAGGTGCTGGAAT 1743
1681 GAGACACACTTACGTTAACTCACTTCAGCCCAACAAAGTGTACCGAGTCCGGATTTCA 1740

1744 GAGACACACTTACGTTAACTCACTTCAGCCAAACAAAGTGTACCGAGTCCGGATTTCA 1803
1741 GCTGGCACTGGCGCTGGCTATGGAGTCCCTTCTCACTGTGATGAGCAACAGCACTGGT 1800
1804 GCTGGCACTGGCGCTGGCTATGGAGTCCCTTCTCACTGTGATGAGCAACAGCACTGGT 1863
1801 GTGCAACAAACAGAGCCATGTTCCCTTTTCCCTCAGAAATTTGAAGTGTGAGGCAAGATG 1860
1864 GTGCAACAAACAGAGCCATGTTCCCTTTTCCCTCAGAAATTTGAAGTGTGAGGCAAGATG 1923
1861 GAGTCCCTGTGTGTGTATGTGAGCGCGCCCTCTCAACCCACCCAGATCTCTGGATACAA 1920
1924 GAGTCCCTGTGTGTGTATGTGAGCGCGCCCTCTCAACCCACCCAGATCTCTGGATACAA 1983
1921 CTCTACTGGGAGAGGTGGGAAACAGAGAGAGAGAGAGTGTGACCGCCCCCGAGGGGT 1980
1984 CTCTACTGGGAGAGGTGGGAAACAGAGAGAGAGAGAGTGTGACCGCCCCCGAGGGGT 2043
1981 CGTGGAGATCAAGCTTGGGACGTGCGGCGCTGCGGCTGGAAGAAAGTGTGAAGCAGTAT 2040
2044 CGTGGAGATCAAGCTTGGGACGTGCGGCGCTGCGGCTGGAAGAAAGTGTGAAGCAGTAT 2103
2041 GAACTGACCCAGTGTAGTCCCTGTGAGCGCGCTACAGAGTGAAGCTCTGTAGCTTTCAACAA 2100
2104 GAACTGACCCAGTGTAGTCCCTGTGAGCGCGCTACAGAGTGAAGCTCTGTAGCTTTCAACAA 2163
2101 CACGAGAGCGCTACGCTGTGTGTGGAAGGGCAAGAGCGAGAGAGGGCCCAACCGCAGAC 2160
2164 CACGAGAGCGCTACGCTGTGTGTGGAAGGGCAAGAGCGAGAGAGGGCCCAACCGCAGAC 2223
2161 CTGCTATTCAGAGGGGGCCACCGCTGCTCTCTGCCCCATGTCCAGCGAGAGTCAACAGC 2220
2224 CTGCTATTCAGAGGGGGCCACCGCTGCTCTCTGCCCCATGTCCAGCGAGAGTCAACAGC 2283
2221 TCCACTTCCATTTGGCTTTCGTTGGAGAAAGCAGACTTTTACACTGTCAAGATGTGCAAC 2280
2284 TCCACTTCCATTTGGCTTTCGTTGGAGAAAGCAGACTTTTACACTGTCAAGATGTGCAAC 2343
2281 TACACTGTGAGCTTTCGCGCCCTGCGGGCTCAGGAATGCTTCCCTGCTCACTATATACC 2340
2344 TACACTGTGAGCTTTCGCGCCCTGCGGGCTCAGGAATGCTTCCCTGCTCACTATATACC 2403
2341 AGCTCTGAGAGAGACATTTCTCATTTGGCGGCTGAGAAACCAATTTACCAAGTGTGCG 2400
2404 AGCTCTGAGAGAGACATTTCTCATTTGGCGGCTGAGAAACCAATTTACCAAGTGTGCG 2463
2401 GTACAGTCCCAACGAGTGTGATATGATGGGCGCTTTTGGCTCGTGTGAGAACGCTCCACC 2460
2464 GTACAGTCCCAACGAGTGTGATATGATGGGCGCTTTTGGCTCGTGTGAGAACGCTCCACC 2523
2461 CTGCTGAGCGGCTTTCAGACCTCTTCTGACCTGCGGCTGAGCGCCCTGACACCATCC 2520
2524 CTGCTGAGCGGCTTTCAGACCTCTTCTGACCTGCGGCTGAGCGCCCTGACACCATCC 2583
2521 ACCGTTTCGTTTACACTGTGTGTCCCGCCACGAGCGCAATGTGTGAGATGTGAGTATCTA 2580
2584 ACCGTTTCGTTTACACTGTGTGTCCCGCCACGAGCGCAATGTGTGAGATGTGAGTATCTA 2643
2581 ATTCTCTACAGCAACCAACCAACCGAGCGCCGAAACCAACAGTGTGACACTGTCTACACAG 2640
2644 ATTCTCTACAGCAACCAACCAACCGAGCGCCGAAACCAACAGTGTGACACTGTCTACACAG 2703
2641 GGAAACATTTCTGAGTGTGAGGTTCATGGCTTGTGAGTGTGACACTGTGAGTATTTCTTCAAG 2700
2704 GGAAACATTTCTGAGTGTGAGGTTCATGGCTTGTGAGTGTGACACTGTGAGTATTTCTTCAAG 2763
2701 ATGGAGCGCCGACAGAGGTGTGGGCGCTTTTCCCGCTTTCAGAGTGTGATTAAT 2760
2764 ATGGAGCGCCGACAGAGGTGTGGGCGCTTTTCCCGCTTTCAGAGTGTGATTAAT 2823
2761 CTGCAAGAGACATTTCTGAGCTCTCTTGGATGTGAC 2796

[illegible]

1321 TCTCTTCACTACCAAAAGGCAAGGGAGTGGACAATGTGGAGTACCAAGTTTGCAGTAAC 1380
1384 TCTCTTCACTACCAAAAGGCAAGGGAGTGGACAATGTGGAGTACCAAGTTTGCAGTAAC 1443
1381 AATGACACACAGAGCTGAGGTTGGGACCTGGAAACCCAAACACAGGATTAATGAGTTCTAC 1440
1444 AATGACACACAGAGCTGAGGTTGGGACCTGGAAACCCAAACACAGGATTAATGAGTTCTAC 1503
1441 GTGGTGGCTACTCCAGCTGGGGCCAGCCGAACTCCAGAGCCAGCCCTGGTGGATACA 1500
1504 GTGGTGGCTACTCCAGCTGGGGCCAGCCGAACTCCAGAGCCAGCCCTGGTGGATACA 1563
1501 CTGGAGCATGTCTCCAGAGCCAGCACCCAGCTTACCTTGTCCAGAGCCCAACCCCTCGGAC 1560
1564 CTGGAGCATGTCTCCAGAGCCAGCACCCAGCTTACCTTGTCCAGAGCCCAACCCCTCGGAC 1623
1561 ATCAGGCTGGATGTGCTCCCTGCTCCAGCTGAGCAATGAGCAAGTGTCTGAGTAC 1620
1624 ATCAGGCTGGATGTGCTCCCTGCTCCAGCTGAGCAATGAGCAAGTGTCTGAGTAC 1683
1621 AAGATAGCTACGTTTGGGAGGAAGATCAGGTTTCTCCAGAGGTGCTGGAAAT 1680
1684 AAGATAGCTACGTTTGGGAGGAAGATCAGGTTTCTCCAGAGGTGCTGGAAAT 1743
1681 GAGACAACTTACGTTAACTCATTGAGCCAAACAAAGTGTACCGAGTCCGGATTCA 1740
1744 GAGACAACTTACGTTAACTCATTGAGCCAAACAAAGTGTACCGAGTCCGGATTCA 1803
1741 GCTGGCACTGGGCTGGCTATGAGTCTCTTCTCAGTGATGAGCAGACACCTGGT 1800
1804 GCTGGCACTGGGCTGGCTATGAGTCTCTTCTCAGTGATGAGCAGACACCTGGT 1863
1801 GTGCACAAACAGAGCATGTTCCCTTGGCCCTGACAGATTCAGAGTGGAGGCAAGATG 1860
1864 GTGCACAAACAGAGCATGTTCCCTTGGCCCTGACAGATTCAGAGTGGAGGCAAGATG 1923
1861 GAGTCCCTGGTGTGATGAGCAGCCGCTCCCTCACCCACACAGATCTCTGATACAAA 1920
1924 GAGTCCCTGGTGTGATGAGCAGCCGCTCCCTCACCCACACAGATCTCTGATACAAA 1983
1921 CTCTACTGGGAGAGTGGGAAACAGAGGAGGAGGAGTGTGACCGCCCCCAGGGGT 1980
1984 CTCTACTGGGAGAGTGGGAAACAGAGGAGGAGGAGTGTGACCGCCCCCAGGGGT 2043
1981 CGTGGAGATCAAGCTTGGGACGTCGGGCCCGCTGAGAGGAGTGAAGCAGTAT 2040
2044 CGTGGAGATCAAGCTTGGGACGTCGGGCCCGCTGAGAGGAGTGAAGCAGTAT 2103
2041 GAACTGACCCAGTTAGTCTCTGAGGCGCTGACAGGCTGAGCTCTGATCTTCAACAAA 2100
2104 GAACTGACCCAGTTAGTCTCTGAGGCGCTGACAGGCTGAGCTCTGATCTTCAACAAA 2163
2101 CACGAGGAGCGCTACGCTGTGTGGAAAGGCAAGCAGGAGGCGCCCAAGCCAGAC 2160
2164 CACGAGGAGCGCTACGCTGTGTGGAAAGGCAAGCAGGAGGCGCCCAAGCCAGAC 2223
2161 CTGCTCTACGAGGGGGCAGCGCTGCTCTGCTCCATGTCCAGAGTGAACAGC 2220
2224 CTGCTCTACGAGGGGGCAGCGCTGCTCTGCTCCATGTCCAGAGTGAACAGC 2283
2221 TCCACTTCCATTTGGCTTGGTGGAAAGAGCCAGACTTTACCACTGTCAAGATTGTCAAC 2280
2284 TCCACTTCCATTTGGCTTGGTGGAAAGAGCCAGACTTTACCACTGTCAAGATTGTCAAC 2343
2281 TACACTGTACGCTTCGGCCCTCGGGGCTCAGGAATGCTTCCCTGGTCACTACTATACC 2340
2344 TACACTGTACGCTTCGGCCCTCGGGGCTCAGGAATGCTTCCCTGGTCACTACTATACC 2403
2341 AGCTCTGGAGAGACATTTCTATTTGGGGCTGAAACCAATTTACCAAGTACGATTGGG 2400
2404 AGCTCTGGAGAGACATTTCTATTTGGGGCTGAAACCAATTTACCAAGTACGATTGGG 2463
2401 GTACAGTCCCAAGGAGTGGATATGGATGGGCCCCCTTTGGCTCGCTGTAGAACCGCTCCACC 2460

2464 GTACAGTCCCAAGGAGTGGATATGGATGGGCCCCCTTTGGCTCGTAGAACGCTCCACC 2523
2461 CTGCTGACCGGCTTCAACACCTCTCTGACCTGGCCCTGAGCCCTGACACCATCC 2520
2524 CTGCTGACCGGCTTCAACACCTCTCTGACCTGGCCCTGAGCCCTGACACCATCC 2583
2521 ACCGTTCTGCTTACACTGGTGTCCCTCCACGAGGCCCCAATGGTGGAGTGTGATCTA 2580
2584 ACCGTTCTGCTTACACTGGTGTCCCTCCACGAGGCCCCAATGGTGGAGTGTGATCTA 2643
2581 ATTCTTACAGACAACACCAACCCAGGCCCCGAAACACAGTGGACACTCTCAACACAG 2640
2644 ATTCTTACAGACAACACCAACCCAGGCCCCGAAACACAGTGGACACTCTCAACACAG 2703
2641 GGAAACATCTTTCAGTGCAGAGTCCATGGCTTGGAGTGCACCTCGGTATTTCTTCAAG 2700
2704 GGAAACATCTTTCAGTGCAGAGTCCATGGCTTGGAGTGCACCTCGGTATTTCTTCAAG 2763
2701 ATGGAGGCCCCGACAGAGTGGGCTTGGGCTCTTTTCCGCTTGCAGATGTGATTACT 2760
2764 ATGGAGGCCCCGACAGAGTGGGCTTGGGCTCTTTTCCGCTTGCAGATGTGATTACT 2823
2761 CTGCAAGAGACATTTCTCAGACTCTCTTGGATGTGCAC 2796
2824 CTGCAAGAGACATTTCTCAGACTCTCTTGGATGTGCAC 2859

RESULT 4
AF176694
LOCUS
DEFINITION Mus musculus neighbor of Punc e11 protein (Nope) mRNA, complete cds.

ACCESSION AF176694
VERSION AF176694.1 GI:7650185
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Salbaum, J.M. and Kappen, C.
TITLE Cloning and expression of nope, a new mouse gene of the immunoglobulin superfamily related to guidance receptors
JOURNAL Genomics 64 (1), 15-23 (2000)
MEDLINE 20175427
PUBMED 10708514

REFERENCE
AUTHORS Salbaum, J.M.
TITLE Direct Submission
JOURNAL Submitted (09-AUG-1999) The Neurosciences Institute, 10640 John Jay Hopkins Drive, San Diego, CA 92121, USA

FEATURES
source
1..6176
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB"
/db_xref="taxon:10090"
/chromosome="9"
/map="band D-E1; 24.59 cR from D9Mit143"
1..6176
/gene="Nope"
1..3759
/gene="Nope"
/note="immunoglobulin superfamily; related to guidance receptors"

gene
CDS
1..3759
/product="neighbor of Punc e11 protein"
/protein_id="AA065930.1"
/db_xref="GI:7650186"

1..6176
/map="band D-E1; 24.59 cR from D9Mit143"
1..6176
/gene="Nope"
1..3759
/gene="Nope"
/note="immunoglobulin superfamily; related to guidance receptors"

1..6176
/map="band D-E1; 24.59 cR from D9Mit143"
1..6176
/gene="Nope"
1..3759
/gene="Nope"
/note="immunoglobulin superfamily; related to guidance receptors"

1..6176
/map="band D-E1; 24.59 cR from D9Mit143"
1..6176
/gene="Nope"
1..3759
/gene="Nope"
/note="immunoglobulin superfamily; related to guidance receptors"

1..6176
/map="band D-E1; 24.59 cR from D9Mit143"
1..6176
/gene="Nope"
1..3759
/gene="Nope"
/note="immunoglobulin superfamily; related to guidance receptors"

PECHTKGLPAPIITWEXDQVTVPEPRLITLIPKWLQILDVQSDAGSYRCVATNSAR
ORFSEASLTVALGSELEATRGQDVIIVAAPENTTVVSGQNVNVECVASADPTFPVSW
VROQKISTDVI VLGRITNLLIASAQPHRSGVYCRANKPLTRDPAFAAABLRVLAAP
ASOAPRALSRTRASTARFVCRASGEPRPALHMDGI PLRPNRGRVQVGGGSLVIT
OIGLQDAGYQCVALENSAGTACAAPLANVVRREGLPSPATRVATPLSSSLVWNER
GALHSEQIIGSLHYQKARGVDNVEYQFVANDTTELQVRLDLEPNTDYEFVVAISOL
FELSTSPALVHTLDDVPSAAPQLTSSPNPSDIRVAMLPLPSSLSNGVLKRYISYG
LGKEDQVFSEVPGNETQLTNSIQPNKXVRVIRISAGTGAGYGVPSQMOHRTPGVNR
QSHVFPAPABELAKVAKMESLIVSWQPPPHPTQISGYKLWYGEVGTBEADGDRPPQGR
GDOAWDVPVRLKKVQYELTQVPCRPVEVKLVA FNKHEDEGVAA WNKTKRKAPTP
DLPIORGPPLPPKHVHAESNSTSINLWRKKPDTTVKLYNYTVRPGMGRNLSLTV
YYSRGGDILLGLKPTKYEFAVQSHGVDMDGPFVSVERSTLDPDRPSPDLRLUS
PLTSTVRLHWCPTPEPGEISVEYLILYSNNHTQPEHMTLLTTEGNI FSAHVHGLS
DTRYFFKMGARTPEYGPFRSLQDVI TLQETFSDSLQVHAWTGI IIVGCLGLCLLAC
MCDVLRSSHREALPLGSSGTPGNPALYTRARLPSPVPAHHELSLVHPRPDQWSP
PPSDVEAKVHSLWGSVSDCRGHSKRKISWAQAGPNWAGSWAGCELPOGSGPRPA
LTRALLPPAGTGQTLILQALVYDGIKNGRKKRSPACRNOVEARVIVHSDPQASKGP
LHLQDLPEPRLFAETLTPSTSGAVDIUSQGDWIGRLBGLGCOPTTSGPERLITCLPEAA
SASCSDLPQPLTAIBRAPGKSCQPKALCPLTVSPSLPRAPVPSAQPV"

misc_feature

1..23

/gene="Nope"

/note="unable to sequence region in mRNA"

BASE COUNT 1316 a 1756 c 1634 g 1470 t

ORIGIN

Query Match 100.0%; Score 2796; DB 10; Length 6176;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2796; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GGGAGCTGCGCAATTTGCCAGGAGCAACTGTCAAGCTGAGCTGTGATGAGGGACCCCTG	60
Db	64	GGGAGCTGCGCAATTTGCCAGGAGCAACTGTCAAGCTGAGCTGTGATGAGGGACCCCTG	123
Qy	61	CAAGTGATCTCGGCGCTGAGCAGGCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG	120
Db	124	CAAGTGATCTCGGCGCTGAGCAGGCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG	183
Qy	121	GCTGGGCTCCGACAGGGTGACATGAGCAAGATGAGACACTGTACTAGACATGAG	180
Db	184	GCTGGGCTCCGACAGGGTGACATGAGCAAGATGAGACACTGTACTAGACATGAG	243
Qy	181	AACCTGACCTGCTACCCAAATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	240
Db	244	AACCTGACCTGCTACCCAAATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	303
Qy	241	AGCGATGATGAGGAAGCTTTAGGATCTGGAAGCTCACTGAGGCGAGCTATTCCTGCTG	300
Db	304	AGCGATGATGAGGAAGCTTTAGGATCTGGAAGCTCACTGAGGCGAGCTATTCCTGCTG	363
Qy	301	GCCACACGCCGCTAGAGGTGTGCGCCAGCAGGTTGCTGTGTGCTCAAGCTTGCACACTC	360
Db	364	GCCACACGCCGCTAGAGGTGTGCGCCAGCAGGTTGCTGTGTGCTCAAGCTTGCACACTC	423
Qy	361	GAAGACTTCTCTGTGCAACCCGAGTCCAGATTGTGAGGAGAAACGGGACAGCAGCTTT	420
Db	424	GAAGACTTCTCTGTGCAACCCGAGTCCAGATTGTGAGGAGAAACGGGACAGCAGCTTT	483
Qy	421	GAATGCCACACCAAGGGCTTCCAGGCCCATCATTAATTCTGGGAAAAGGACAGGTAAC	480
Db	484	GAATGCCACACCAAGGGCTTCCAGGCCCATCATTAATTCTGGGAAAAGGACAGGTAAC	543
Qy	481	GTGCTGAGGAGCCCGGCTCATCTCTCCCAAGTGGCTCTCCAGATCCTAGATGTC	540
Db	544	GTGCTGAGGAGCCCGGCTCATCTCTCCCAAGTGGCTCTCCAGATCCTAGATGTC	603
Qy	541	CAGGACAGTGATCAGGCTCTTACCGCTGCTGCGCCACCAATTCAGCCCGCCCAACGATTC	600
Db	604	CAGGACAGTGATCAGGCTCTTACCGCTGCTGCGCCACCAATTCAGCCCGCCCAACGATTC	663
Qy	601	AGCCAGGAGGCTCGCTCACTGTGGCCCTCAGAGGGCTTTTGGAGGCTTACAGGGGGCAG	660
Db	664	AGCCAGGAGGCTCGCTCACTGTGGCCCTCAGAGGGCTTTTGGAGGCTTACAGGGGGCAG	723

Qy	661	GATGTGTCTCATTTGTGGCAGCCCCCAGAGAAACCAACCGTGTGTCTGGACAGAAATGTAGTG	720
Db	724	GATGTGTCTCATTTGTGGCAGCCCCCAGAGAAACCAACCGTGTGTCTGGACAGAAATGTAGTG	783
Qy	721	ATGGAGTGCCTGTGCTCTGTCTGACCCACCCCTTTTGTGTCTCTGGTCCGACAGGATGGA	780
Db	784	ATGGAGTGCCTGTGCTCTGTCTGACCCACCCCTTTTGTGTCTCTGGTCCGACAGGATGGA	843
Qy	781	AAGCTATCTCCACGGATGTCACTGTCTGGGGCGGACCAATCTACTACTCATCGCCAGCGCG	840
Db	844	AAGCTATCTCCACGGATGTCACTGTCTGGGGCGGACCAATCTACTACTCATCGCCAGCGCG	903
Qy	841	CAGCTCTCGGCACTCTGGAGTCTATGTCTGCGGAGCCAAAGCCCTCTCAGCGGTGACTTC	900
Db	904	CAGCTCTCGGCACTCTGGAGTCTATGTCTGCGGAGCCAAAGCCCTCTCAGCGGTGACTTC	963
Qy	901	GCACCTGCGGCTGTGAGCTCCGAGTCTTGTCTGCCACGACATCTGCGAGGCAACCGGAG	960
Db	964	GCACCTGCGGCTGTGAGCTCCGAGTCTTGTCTGCCACGACATCTGCGAGGCAACCGGAG	1023
Qy	961	GCGCTCTCGGCGAGCGGGGCGCAGCACCGCGCTTGTGTGTGCGGGGCTCCGGGGAGCCA	1020
Db	1024	GCGCTCTCGGCGAGCGGGGCGCAGCACCGCGCTTGTGTGTGCGGGGCTCCGGGGAGCCA	1083
Qy	1021	CGGCTCTCGGCTGTGCACTGGCTGTGCAAGCGGATCCCGTTGTGCGACCCCAATGGGCGGCTCAAG	1080
Db	1084	CGGCTCTCGGCTGTGCACTGGCTGTGCAAGCGGATCCCGTTGTGCGACCCCAATGGGCGGCTCAAG	1143
Qy	1081	GTGACGGGCGGTGTGCGGCTGTGCTGTATCTCAGATGTGCGCTGTGAGGACGCTGGCTAC	1140
Db	1144	GTGACGGGCGGTGTGCGGCTGTGCTGTATCTCAGATGTGCGCTGTGAGGACGCTGGCTAC	1203
Qy	1141	TACAGTGCCTGTGACGAGAAACAGCGCGGAACTGCTGCTGTGCGCTGTGCGCTGTGCGGTA	1200
Db	1204	TACAGTGCCTGTGACGAGAAACAGCGCGGAACTGCTGCTGTGCGCTGTGCGCTGTGCGGTA	1263
Qy	1201	GTGCTGTGCGGAGGGCTGTCCCGAGCGCCCGACTCGGGTCAAGCACCGCGCTCTGAGCAGC	1260
Db	1264	GTGCTGTGCGGAGGGCTGTCCCGAGCGCCCGACTCGGGTCAAGCACCGCGCTCTGAGCAGC	1323
Qy	1261	TCTCTGTGTGTGTGCTGTGCGGCGGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG	1320
Db	1324	TCTCTGTGTGTGTGCTGTGCGGCGGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG	1383
Qy	1321	TCTCTTCACTACCAAAAGGCAAGGGGAGTGACAAATGTGAGTACCAAGTTTGCAGTAAC	1380
Db	1384	TCTCTTCACTACCAAAAGGCAAGGGGAGTGACAAATGTGAGTACCAAGTTTGCAGTAAC	1443
Qy	1381	AATGACACACAGAGCTGTGAGTTTGGGACCTGGAACTTGAACCCCAACAGGATTTAGTCTAC	1440
Db	1444	AATGACACACAGAGCTGTGAGTTTGGGACCTGGAACTTGAACCCCAACAGGATTTAGTCTAC	1503
Qy	1441	GTGCTGCGCTTACTCCAGCTGGGGGCGAGCGGAACTTCCAGCCCGAGCTGTGCTGATACA	1500
Db	1504	GTGCTGCGCTTACTCCAGCTGGGGGCGAGCGGAACTTCCAGCCCGAGCTGTGCTGATACA	1563
Qy	1501	CTGACGATCTCCCGAGCGCAGCACCCAGCTTACCTTGTCTCCAGCCCGACCCCTCGGAC	1560
Db	1564	CTGACGATCTCCCGAGCGCAGCACCCAGCTTACCTTGTCTCCAGCCCGACCCCTCGGAC	1623
Qy	1561	ATCAGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1620
Db	1624	ATCAGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1683
Qy	1621	AAGTAGAGTACGGTTTGGGAAAGGAGATCAGGTTTCTCCACCGAGGTCCTGGAAAT	1680
Db	1684	AAGTAGAGTACGGTTTGGGAAAGGAGATCAGGTTTCTCCACCGAGGTCCTGGAAAT	1743
Qy	1681	GAGACACAACTTACGTTTAAACTCACTTTCAGGCAACAAAGTGTACCGAGTCCGGAATTC	1740
Db	1744	GAGACACAACTTACGTTTAAACTCACTTTCAGGCAACAAAGTGTACCGAGTCCGGAATTC	1803
Qy	1741	GCTGGCACTGGCGCTGTGCTATGGAGTCCCTTCTCAGTGTGATGACAGCAGGACACCTGGT	1800

Qy 1 GGGAGCTGCCATTGGCCCGAGGAGCAAACTGTCTAAGCTGTAGCTGTGTATGAGGGACCCCTG 60
Db 113 GGGAGCTGCCATTGGCCCGAGGAGCAAACTGTCTAAGCTGTAGCTGTGTATGAGGGACCCCTG 172
Qy 61 CAAGTATCTTGGGCCCTGAGCAGGCTGTGTGTCTGGACTGTGCACTTTTGGGGGCTTACAGCT 120
Db 173 CAAGTATCTTGGGCCCTGAGCAGGCTGTGTGTCTGGACTGTGCACTTTTGGGGGCTTACAGCT 232
Qy 121 GCTGGGCTCCGACCGAGGTGACATGAGCAGCAAGATGAGACACATGTACTAGAGCATGAG 180
Db 233 GCTGGGCTCCGACCGAGGTGACATGAGCAGCAAGATGAGACACATGTACTAGAGCATGAG 292
Qy 181 AACCTGACCTGTCTACCAATGGCTCCCTGTGGCTGTCTCACCCCTAGAGCAAGAGAC 240
Db 293 AACCTGACCTGTCTACCAATGGCTCCCTGTGGCTGTCTCACCCCTAGAGCAAGAGAC 352
Qy 241 AGCGATGATGAGGAGCTCTTAGATCTGGAAGGTCTCATGAGGCGACGTATTCCTGTCTG 300
Db 353 AGCGATGATGAGGAGCTCTTAGATCTGGAAGGTCTCATGAGGCGACGTATTCCTGTCTG 412
Qy 301 GCCCAGCCGCTTAGGAGTGTGGCCAGCCAGGCTGTGTGTCTGAGTCTGAGCTTGCACACTC 360
Db 413 GCCCAGCCGCTTAGGAGTGTGGCCAGCCAGGCTGTGTGTCTGAGTCTGAGCTTGCACACTC 472
Qy 361 GAAGACTTCTCTCTGCAACCGAGTCCAGATTTGGAGGAGAACCGGACAGCACTGCTTT 420
Db 473 GAAGACTTCTCTCTGCAACCGAGTCCAGATTTGGAGGAGAACCGGACAGCACTGCTTT 532
Qy 421 GAATGCCACCAAGGCGCTTCCAGCCCCCATCATTTCTTGGGAAAGGACCAAGTGAAC 480
Db 533 GAATGCCACCAAGGCGCTTCCAGCCCCCATCATTTCTTGGGAAAGGACCAAGTGAAC 592
Qy 481 GTGGCTGAGGAGCCCCGGCTCATCACTTTCCTCAAGTGGCTCTCCAGATCCTAGATGTC 540
Db 593 GTGGCTGAGGAGCCCCGGCTCATCACTTTCCTCAAGTGGCTCTCCAGATCCTAGATGTC 652
Qy 541 CAGGACAGTATGATGAGGCTCTTACCGCTGTGGCCACCAATTCAGCCCGCCCAACGATTC 600
Db 653 CAGGACAGTATGATGAGGCTCTTACCGCTGTGGCCACCAATTCAGCCCGCCCAACGATTC 712
Qy 601 AGCCAGGAGGCTCGCTCATCTGTGGCCCTCAGAGGGTCTTTGAGGCTTACAGGGGGCAG 660
Db 713 AGCCAGGAGGCTCGCTCATCTGTGGCCCTCAGAGGGTCTTTGAGGCTTACAGGGGGCAG 772
Qy 661 GATGTGTCTATTGGGAGCCCGCAGAGAACCAACAGGTAGTGTCTGACAGATGTTAGTG 720
Db 773 GATGTGTCTATTGGGAGCCCGCAGAGAACCAACAGGTAGTGTCTGACAGATGTTAGTG 832
Qy 721 ATGAGTGTGGGCTCTGTGTGACCCCAACCCCTTTTGTGTCTGGGTCCGACAGATGGA 780
Db 833 ATGAGTGTGGGCTCTGTGTGACCCCAACCCCTTTTGTGTCTGGGTCCGACAGATGGA 892
Qy 781 AAGCCTATCTCCAGGATGTCATCGTTCTGGCCCGGACCAATCTACTCATCGCCAGCGG 840
Db 893 AAGCCTATCTCCAGGATGTCATCGTTCTGGCCCGGACCAATCTACTCATCGCCAGCGG 952
Qy 841 CAGCCTCGGCACTGTGAGTCTATGTCTGCGAGCCAAAGAGCCCTCATCGGTGACTTC 900
Db 953 CAGCCTCGGCACTGTGAGTCTATGTCTGCGAGCCAAAGAGCCCTCATCGGTGACTTC 1012
Qy 901 GCCATCGGCTGTGAGTCTCGAGTCTTGTCTGCCCGACGATCTCGAGGACCCGAG 960
Db 1013 GCCATCGGCTGTGAGTCTCGAGTCTTGTCTGCCCGACGATCTCGAGGACCCGAG 1072
Qy 961 GCGCTCTCGGAGACCGGGCCAGCACCGCGCTTCTGTGTGCGGGGCTCGGGGAGCCA 1020
Db 1073 GCGCTCTCGGAGACCGGGCCAGCACCGCGCTTCTGTGTGCGGGGCTCGGGGAGCCA 1132
Qy 1021 CGGCCCGGCTGCACTGTGCTGACGAGCGGATCCCGTTGTGAGACCAATGGGCGGCTCAAG 1080
Db 1133 CGGCCCGGCTGCACTGTGCTGACGAGCGGATCCCGTTGTGAGACCAATGGGCGGCTCAAG 1192

Qy 1081 GTGCAAGGCGGCTGGCGGCAAGCTTGGTCACTCATCTAGATCGGCTGTGAGGACGCTGGCTAC 1140
Db 1193 GTGCAAGGCGGCTGGCGGCAAGCTTGGTCACTCATCTAGATCGGCTGTGAGGACGCTGGCTAC 1252
Qy 1141 TACCACTGTGTAGCAGAAACACGCGGGGAACTGCTGTGTGCGGCTGTGCGGCTTGGGGTA 1200
Db 1253 TACCACTGTGTAGCAGAAACACGCGGGGAACTGCTGTGTGCGGCTGTGCGGCTTGGGGTA 1312
Qy 1201 GTGGTGTGCGGAGGGGCTGCCAGCGCCCGACTCTGGGTCAACGACCGCGCTGTGAGCAGC 1260
Db 1313 GTGGTGTGCGGAGGGGCTGCCAGCGCCCGACTCTGGGTCAACGACCGCGCTGTGAGCAGC 1372
Qy 1261 TCTCTGTGTGTGTGGCTGTGGAGCGGCTGTGATGTGCAAGCAGCAAAATCATTTGGCTTC 1320
Db 1373 TCTCTGTGTGTGTGGCTGTGGAGCGGCTGTGATGTGCAACGAGCAAAATCATTTGGCTTC 1432
Qy 1321 TCTCTTCACTTACCAAAAGGCAAGGGGAGTGGACAAATGTGGAGTACCAATTTGCAATTAAC 1380
Db 1433 TCTCTTCACTTACCAAAAGGCAAGGGGAGTGGACAAATGTGGAGTACCAATTTGCAATTAAC 1492
Qy 1381 AATGACACCAACAGAGCTGCAAGGTTCCGGACTCTGGAACCCAAACAGGATTAATGATTTCTAC 1440
Db 1493 AATGACACCAACAGAGCTGCAAGGTTCCGGACTCTGGAACCCAAACAGGATTAATGATTTCTAC 1552
Qy 1441 GTGGTGGCTTACTCCAGCTGTGGGGCCAGCCGAACTCTCAGCCAGCCCTGTGTGCTATCA 1500
Db 1553 GTGGTGGCTTACTCCAGCTGTGGGGCCAGCCGAACTCTCAGCCAGCCCTGTGTGCTATCA 1612
Qy 1501 CTGGAAGATGTCTCCAGCGGAGCAGCCAGCTTACTTGTCTCAGCCCAACCCCTCGGAC 1560
Db 1613 CTGGAAGATGTCTCCAGCGGAGCAGCCAGCTTACTTGTCTCAGCCCAACCCCTCGGAC 1672
Qy 1561 ATCAGGTTGTCATGGCTGTGCCCTCTCCAGCTGTGAGCAATGGAACAGGCTGTGAGTAC 1620
Db 1673 ATCAGGTTGTCATGGCTGTGCCCTCTCCAGCTGTGAGCAATGGAACAGGCTGTGAGTAC 1732
Qy 1621 AAGATAGATGACGTTTGGGAAAGGAGATCAGGTTTCTTCCACCGAGGTGCTTGGAAAT 1680
Db 1733 AAGATAGATGACGTTTGGGAAAGGAGATCAGGTTTCTTCCACCGAGGTGCTTGGAAAT 1792
Qy 1681 GAGACAACTTACGTTAAACTCATCTTACGCAAAACAAAGTGTACCGAGTCCGGAATTTCA 1740
Db 1793 GAGACAACTTACGTTAAACTCATCTTACGCAAAACAAAGTGTACCGAGTCCGGAATTTCA 1852
Qy 1741 GCTGGCACTGGCGCTGCTATGAGTCCCTTCTCAGTGTGATGAGCAGACACCTGCT 1800
Db 1853 GCTGGCACTGGCGCTGCTATGAGTCCCTTCTCAGTGTGATGAGCAGACACCTGCT 1912
Qy 1801 GTGCAAAACAGAGCCATGTTCCCTTGTGCCCTGTGCAAAATGGAAGGTGAGGGCAAGATG 1860
Db 1913 GTGCAAAACAGAGCCATGTTCCCTTGTGCCCTGTGCAAAATGGAAGGTGAGGGCAAGATG 1972
Qy 1861 GAGTCCCTGTGTGTCTATGGAAGCGGCTTCTTCAACCCCAACCGATCTCTGTGATACAA 1920
Db 1973 GAGTCCCTGTGTGTCTATGGAAGCGGCTTCTTCAACCCCAACCGATCTCTGTGATACAA 2032
Qy 1921 CTCTACTGGGAGAGGTGGGAAACAGAGGAGGAGGAGTGTGATGATGATGATGATGATGAT 1980
Db 2033 CTCTACTGGGAGAGGTGGGAAACAGAGGAGGAGGAGTGTGATGATGATGATGATGATGAT 2092
Qy 1981 CGTGGAGATCAAGCTTGGGACGTCTGGGCGGCTGGCTGGAAGAAAGTGAAGCAGTAT 2040
Db 2093 CGTGGAGATCAAGCTTGGGACGTCTGGGCGGCTGGCTGGAAGAAAGTGAAGCAGTAT 2152
Qy 2041 GAACTGACCCAGTGTCTCTGGAAGCGGCTTACGAGGTGAAGCTCTGTAGCTTTCAACAA 2100
Db 2153 GAACTGACCCAGTGTCTCTGGAAGCGGCTTACGAGGTGAAGCTCTGTAGCTTTCAACAA 2212
Qy 2101 CAGAGGACCGGCTTACGCTGTGTGTGGAAGGGCAACGAGGAGGCGGCGCCACCGCAGAC 2160
Db 2213 CAGAGGACCGGCTTACGCTGTGTGTGGAAGGGCAACGAGGAGGCGGCGCCACCGCAGAC 2272
Qy 2161 CTGCTTATCAGAGGGGGCCACCGCTGCTCTGCTCCCTCTGTCACGACGAGTCAACACAGC 2220

Db	2273	CTGCTATTCAGAGGGGGCCACCGCTGCTCTCTGCCCCATGTGCA CGAGAGTCAACAGC	2332
Qy	2221	TCCACTTCCATTTGGCTTTGGTGGAAGACGAGACTTTACCACTGTGCAAGATTGTCAAC	2280
Db	2333	TCCACTTCCATTTGGCTTTGGTGGAAGACGAGACTTTACCACTGTGCAAGATTGTCAAC	2392
Qy	2281	TACACTGTACGCTTGGCCCCCTGGGGCTCAGGAATGCTTCCCTGGTCACCTACTATACC	2340
Db	2393	TACACTGTACGCTTGGCCCCCTGGGGCTCAGGAATGCTTCCCTGGTCACCTACTATACC	2452
Qy	2341	AGCTCTGGAGAAGACATTCTCATTTGGCGGCTGAAACCAATTACCAAGTACGAGTTTGC	2400
Db	2453	AGCTCTGGAGAGACATTCTCATTTGGCGGCTGAAACCAATTACCAAGTACGAGTTTGC	2512
Qy	2401	GTACAGTCCACGAGTGGATATGGATGGGCCCCCTTTGGCTCGTGTAGAAACGCTCCACC	2460
Db	2513	GTACAGTCCACGAGTGGATATGGATGGGCCCCCTTTGGCTCGTGTAGAAACGCTCCACC	2572
Qy	2461	CTGCTGACCGGCTTCAAACACTCTCTCTGACTGCGGCTTGAGCCCTCTGACACCATCC	2520
Db	2573	CTGCTGACCGGCTTCAAACACTCTCTCTGACTGCGGCTTGAGCCCTCTGACACCATCC	2632
Qy	2521	ACGGTTCGGTTACACTGGTGTCCCCCACA CGGAGCCCAATGGTGAGATTGTGGAGTATCTA	2580
Db	2633	ACGGTTCGGTTACACTGGTGTCCCCCACA CGGAGCCCAATGGTGAGATTGTGGAGTATCTA	2692
Qy	2581	ATTCTCTACAGCAACACACACCCAGCCGAA CACAGTGGACACTGCTCACCACAGAG	2640
Db	2693	ATTCTCTACAGCAACACACACCCAGCCGAA CACAGTGGACACTGCTCACCACAGAG	2752
Qy	2641	GGAAACATCTTCAGTGCAGAGGTCATGGCTTAGAGATGCACACTCGGTATTTCTCAAG	2700
Db	2753	GGAAACATCTTCAGTGCAGAGGTCATGGCTTAGAGATGCACACTCGGTATTTCTCAAG	2812
Qy	2701	ATGGAGCCCGCACAGAGGTGGGCCCCCTTTCCGCTTGCAGGATGTGATTACT	2760
Db	2813	ATGGAGCCCGCACAGAGGTGGGCCCCCTTTCCGCTTGCAGGATGTGATTACT	2872
Qy	2761	CTGCAAGAGACATTCTCAGACTCTCTGGATGTGCAC	2796
Db	2873	CTGCAAAAGACATTCTCAGACTCTCTGGATGTGCAC	2908

RESULT	6
LOCUS	AB052621 Mus musculus DDM36E mRNA, complete cds.
DEFINITION	6222 bp linear
ACCESSION	AB052621
VERSION	AB052621.1 GI:11862940
KEYWORDS	Mus musculus (house mouse)
SOURCE	Mus musculus
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
REFERENCE	1 (sites) Murakami,H., Nakamata,T., Nakayama,T., Yamamoto,H., Hosaka,T., Aoyama,T., Nagayama,S.; Oka,M., Kiyono,T., Sasaki,M.S., Nakamura,T. and Toguchida,J. Up-regulation of a ras effector and down-regulation of a cell adhesion molecule are associated with transformation of osteoblasts
AUTHORS	Unpublished
TITLE	2 (bases 1 to 6222) Toguchida,J., Nakamata,T., Murakami,H., Nakayama,T. and Nakamura,T.
JOURNAL	Direct Submission
REFERENCE	Submitted (11-DEC-2000) Junya Toguchida, Kyoto University, Institute for Frontier Medical Sciences; 53 Kawahara-cho, Shogoin, Sakyoku-ku, Kyoto, Kyoto 606-8507, Japan (E-mail:togiun@frontier.kyoto-u.ac.jp, Tel:81-75-751-4134, Fax:81-75-751-4144)
FEATURES	Location/Qualifiers 1..6222 /organism="Mus musculus"
SOURCE	

```

/mol_type="mrna"
/strain="C57BL/6BA"
/db_xref="taxon:10090"
/clone="DDM368"
/cell_line="MWC2"
/cell_type="osteoblast"
/notes="similar to Nope (NM_020043)-vector:pDDM368ORP"
1. 6222
/gene="DDM368"
50..3811
/gene="DDM368"
/notes="similar to Nope (NP_064427)"
/codon_start=1
/evidence=experimental
/product="DDM368"
/protein_id="BAB19279.1"
/db_xref="GI:11862941"
/tranlation="MARADTGRLGLVLTFCLLSARGELPLPQSTTVKLSCEGPLOVIV
LQPEARVLDCITGATAGPPTRVWSKGDQTVLEHNLHLLPNSIMLSSPLEQDS
DDEALRIWKVTEGSSCIAHSPFLVWASQVAVVKLATIEDFLHPSIMQVLENGTAR
PFCHTGGLPAPITWKEQDQVTPPEBSRLITLPGNVLIQILDVQSDAGSVYCNATGAR
QRFSQEASITVALRGSLAQGVVIAAFTPTVSGQSVMECAGVADPVPFVSW
VROQKPISTDIVILGRTNLLTASAPRHSVVVCRANKPRTDRDPATAAELRLAALP
ATSQAPEARLSRTASTARPCVRSAGEPRPALHWHGKIPLRPNGRVKVQGGGSLVIT
QIGLDAGTYQCVAENSAGTACAAPLAVVVBEGLSAPATRVATATPLSSSVLVAMER
PELHSEIIGSLHYQKQANVQYQFANNDDTTLQVRDLDEPNTDIIFYVVAISQL
GASRTSSPALVTLQDDVPSAAQPTLLSSPNPDSIRVAMLPPLSSLSNGQVLYKIEYG
LGBEBOPEFTVGNKEQTLLNSLPKNKYVRVRIAGTCAGYGVFSQWQHRTPTGVH
NOSHVPFAEELKAMKSLVVSWMOPPHPTQISGVLYWREVGVEERADGPPPGG
RDQANDVGPVRLKKVKQYBELTQLVPELRYEVLVAPNKHEDGYAANVMKTEKAPT
PDLPQRGPELPPAHVHASENSISLWRMKPDPFTVKIIVNTIVTFPGPMLNASHLV
TYTSGEEDILIGLKLPTFKTSYFVQSHGVDMDGPFVSVERSTLPDRPSTPPSDRLI
SPLTSTVRLHRCPTPENGEIVEYLILYNNHTQPEHOWTLTTGNIISAEVHGLE
SDTRVFPKMGARTVEGPFRLQDVIITLQKTFSDSLDHAVTGI17VCGLICLILA
CMCAGLRSSHREALPGLSSGTPGNPALYTRALRGPPSPVAHLELSLHVPRPQDMSI
PPPSVEDVGAHVHSLMGSVSDCRGSHKKI1SWAQQGPNWAGSWAGCEL1PQSGSPRP
ALTRLHPAGTGTLQULLQALVTDATKSNRGKPPSPACRNQVEARVIVHSDFGASKRG
POLHLQDLPEBPLTAEPLTSPGAVDLSQGDMLGRELGGCQPTTSGSPRLTCLPRA
ASAGSCSDLPQSTAT1EAPAGSKSCPKALCPILTVPSPLPRAPVSSAQVP"
BASE COUNT      1321 a 1772 c 1653 g 1476 t
ORIGIN

Query Match      98.8%; Score 2762.2; DB 10; Length 6222;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2783; Conservative 0; Mismatches 13; Indels 3; Gaps 1;

Qy  1  GGGGAGCTGCATTGGCCCGAGGAGACAACCTGTCAAGCTGAGCTGTGATGAGGAGACCCCTG 60
Db 113 GGGGAGCTGCATTGGCCCGAGGAGACAACCTGTCAAGCTGAGCTGTGATGAGGAGACCCCTG 172

Qy  61  CAAGTGATCTGGGGCCCTGAGCAGGCTGTGGTGTCTGACTGCACCTTTGGGGGCTACAGCT 120
Db 173 CAAGTGATCTGGGGCCCTGAGCAGGCTGTGGTGTCTGACTGCACCTTTGGGGGCTACAGCT 232

Qy 121  GCTGGGCTCCGACGAGGTGA CATGAGCAAGGATGAGACACTGTACTAGACGATGAG 180
Db 233 GCTGGGCTCCGACGAGGTGA CATGAGCAAGGATGAGACACTGTACTAGACGATGAG 292

Qy 181  AACCTGCACCTGCTACCCAAATGCTCCCTGTGCTGCTCTCAACCCCTAGAGCAGAGAC 240
Db 293 AACCTGCACCTGCTACCCAAATGCTCCCTGTGCTGCTCTCAACCCCTAGAGCAGAGAC 352

Qy 241  AGCGATGATGAGGAGCTCTTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTCTGCTG 300
Db 353 AGCGATGATGAGGAGCTCTTAGGATCTGGAAGGTCACTGAGGGCAGCTATTCTCTGCTG 412

Qy 301  GCCCACAAGCCCGCTAGGAGTGGTGGCCAGCCAGGTTCGTGTGGTCAAGCTTGGCCACATC 360
Db 413 GCCCACAAGCCCGCTAGGAGTGGTGGCCAGCCAGGTTCGTGTGGTCAAGCTTGGCCACATC 472

Qy 361  GAAGACTTCTCTGTGACCCCGAGTCCAGATTTGTGAGGAGGAACGGGACAGGACGCTTT 420
Db 473 GAAGACTTCTCTGTGACCCCGAGTCCAGATTTGTGAGGAGGAACGGGACAGGACGCTTT 532

```

Qy	421	GAATGCCACCAAGGGCTTCCAGCCCCCATCTATTCTTGGGAAAGGACCAAGTGACC	480
Db	533	GAATGCCACCAAGGGCTTCCAGCCCCCATCTATTCTTGGGAAAGGACCAAGTGACC	592
Qy	481	GTGCTGAGGAGCCCGGCTCATCTCTTCCAGAGTGGCTCTCCAGATCCTAGATGTC	540
Db	593	GTGCTGAGGAGTCCCGGCTCATCTCTTCCAGAGTGGCTCTCCAGATCCTAGATGTC	652
Qy	541	CAGGACAGTGAATCAGGCTCTCTACCGCTGCGTGCACCAATTCAGCGCCGCAACGATTC	600
Db	653	CAGGACAGTGAATCAGGCTCTCTACCGCTGCGTGCACCAATTCAGCGCCGCAACGATTC	712
Qy	601	AGCCAGAGGCTCTCGCTCATCTGTGGCTCTCAGAGGCTCTTTGGAGGCTACAGAGGGGAG	660
Db	713	AGCCAGAGGCTCTCGCTCATCTGTGGCTCTCAGAGGCTCTTTGGAGGCTACAGAGGGGAG	772
Qy	661	GATGTGTCTATTGGGAGCCGCCAGAGAACACACAGTAGTGTCTGGACAGATGTAGTG	720
Db	773	GATGTGTCTATTGGGAGCCGCCAGAGAACACACAGTAGTGTCTGGACAGATGTAGTG	832
Qy	721	ATGAGTGTGCTGGCTCTGCTGACCCACCCCTTTTGTGCTCGGTCGACAGATGGA	780
Db	833	ATGAGTGTGCTGGCTCTGCTGACCCACCCCTTTTGTGCTCGGTCGACAGATGGA	892
Qy	781	AAGCCTATCTCCAGGATGTCTATGTTCTGGGCGGACCAATCTACTCATCGCCAGCGG	840
Db	893	AAGCCTATCTCCAGGATGTCTATGTTCTGGGCGGACCAATCTACTCATCGCCAGCGG	952
Qy	841	CAGCTCTGGCACTCTGGAGTCTATGTCTGGCGAGCCAAAGAGCCCTCAGCGTGACTTC	900
Db	953	CAGCTCTGGCACTCTGGAGTCTATGTCTGGCGAGCCAAAGAGCCCTCAGCGTGATTC	1012
Qy	901	GCACCTGGCTGCTGAGTCCGAGTCTGCTGCCCGCCAGCATCTCGAGGCAACCGAG	960
Db	1013	GCACCTGGCTGCTGAGTCTGAGTCTGCTGCCCGCCAGCATCTCGAGGCAACCGAG	1072
Qy	961	CGCTCTCTCGGAGCGCGGGCCAGACACCGCGCTTCTGTGTCGGGGGTCCGGGGAGCCA	1020
Db	1073	CGCTCTCTCGGAGCGCGGGCCAGACACCGCGCTTCTGTGTCGGGGGTCCGGGGAGCCA	1132
Qy	1021	CGGCTCGGCTGCACTGGCTGCAAGAGGATCCGTTGCGAACCAATGGGCGCGTCAAG	1080
Db	1133	CGGCTCGGCTGCACTGGCTGCAAGAGGATCCGTTGCGAACCAATGGGCGCGTCAAG	1192
Qy	1081	GTGCAGGCGGTGGGCGAGTCTGGTCACTCATCTCAGATCGGCTGCGAGCGCTGGCTAC	1140
Db	1193	GTGCAGGCGGTGGGCGAGTCTGGTCACTCATCTCAGATCGGCTGCGAGCGCTGGCTAC	1252
Qy	1141	TACCACTGCTAGCAGAAACAGCGCGGGAATGCTGCTGTGCGCTGCGCCCTGGCGGTA	1200
Db	1253	TACCACTGCTAGCAGAAACAGCGCGGGAATGCTGCTGTGCGCTGCGCCCTGGCGGTA	1312
Qy	1201	GTGTGCGGAGGGCTGCCCCAGCGCCCGACTCGGGTCAAGCCACCGCCCTGAGCAGC	1260
Db	1313	GTGTGCGGAGGGCTGCCCCAGCGCCCGACTCGGGTCAAGCCACCGCCCTGAGCAGC	1372
Qy	1261	TCCTCTGTGCTGTGGCTTGGAGCGGCTGAGTGTGACAGCGAGCAATCAITGGGCTTC	1320
Db	1373	TCCTCTGTGCTGTGGCTTGGAGCGGCTGAGTGTGACAGCGAGCAATCAITGGGCTTC	1432
Qy	1321	TCTCTTCACTAACAAAGGCAAGGGAGTGACAAATGTGGAGTACCACTTTGCAAGTAAAC	1380
Db	1433	TCTCTTCACTAACAAAGGCAAGGGAGTGACAAATGTGGAGTACCACTTTGCAAGTAAAC	1492
Qy	1381	AATGACACCAAGAGCTGACGTTTGGGACCTGGAAACCAACAGCGATTTAGTCTTAC	1440
Db	1493	AATGACACCAAGAGCTGACGTTTGGGACCTGGAAACCAACAGCGATTTAGTCTTAC	1552
Qy	1441	GTGTGGGCTTACTCCAGCTGGGGGCGAGCGAACTCTCAGCCAGCGCTGGTGATACA	1500
Db	1553	GTGTGGGCTTACTCCAGCTGGGGGCGAGCGAACTCTCAGCCAGCGCTGGTGATACA	1612

Qy	1501	CTGAGCAGATGTCCTCCAGCGCAGCACCCAGCTTACTCTTGTCCAGCCCCCAACCCCTCGGAC	1560
Db	1613	CTGAGCAGATGTCCTCCAGCGCAGCACCCAGCTTACTCTTGTCCAGCCCCCAACCCCTCGGAC	1672
Qy	1561	ATCAGGGTGCATGGCTGCGCTGCTCCAGCTGAGCAATGGAAGAGTGTCTGAGATAC	1620
Db	1673	ATCAGGGTGCATGGCTGCGCTGCTCCAGCTGAGCAATGGAAGAGTGTCTGAGATAC	1732
Qy	1621	AAGATAGATGACGGTTTGGGGAAG--GAAGATCAGGTTTCTCCACCGAGGTGCCTGGA	1677
Db	1733	AAGATAGATGACGGTTTGGGGAAGAGATCAGGTTTCTCCACCGAGGTGCCTGGA	1792
Qy	1678	AATGAGACACAATTAAGTTAAATCACTTTCAGCCAAAGAGTGTACCGAGTCCGATTT	1737
Db	1793	AATGAGACACAATTAAGTTAAATCACTTTCAGCCAAAGAGTGTACCGAGTCCGATTT	1852
Qy	1738	TCAGCTGGCACTGCGCTGCTGATGAGTCCCTCTCAGTGGATGCAAGCAGGACACCT	1797
Db	1853	TCAGCTGGCACTGCGCTGCTGATGAGTCCCTCTCAGTGGATGCAAGCAGGACACCT	1912
Qy	1798	GGTGTGCAACAACAGAGCCATGTTCCCTTTGGCCCTGACAGAAATGAAGGTGAGGGCAAG	1857
Db	1913	GGTGTGCAACAACAGAGCCATGTTCCCTTTGGCCCTGACAGAAATGAAGGTGAGGGCAAG	1972
Qy	1858	ATGAGTCTCTGCTGCTGATGAGTCCCTCTCAGCCCGCTCAGCCCAAGATCTCTGAGTAC	1917
Db	1973	ATGAGTCTCTGCTGCTGATGAGTCCCTCTCAGCCCGCTCAGCCCAAGATCTCTGAGTAC	2032
Qy	1918	AAACTCTACTTGGGAGAGGTGGGAAACAGAGAGGAGGAGAGTGTGACCGCCCCCAGGG	1977
Db	2033	AAACTCTACTTGGGAGAGGTGGGAAACAGAGAGGAGGAGAGTGTGACCGCCCCCAGGG	2092
Qy	1978	GGTGTGAGATCAAGCTTGGGAGCTGCGGCGCTGAGGAGGAGGAGTGTGAGAGGAGAGCAG	2037
Db	2093	GGTGTGAGATCAAGCTTGGGAGCTGCGGCGCTGAGGAGGAGGAGTGTGAGAGGAGAGCAG	2152
Qy	2038	TATGAACTGACCCAGTTAGTCCCTTGGCAGGCGGTACGAGGTGAGCTCTGAGCTTTCAAC	2097
Db	2153	TATGAACTGACCCAGTTAGTCCCTTGGCAGGCGGTACGAGGTGAGCTCTGAGCTTTCAAC	2212
Qy	2098	AAACAAGAGAGCGGTACGCTGCTGTGTGGAAGGGCAAGACGGAGAGGGCCCAAGCCCA	2157
Db	2213	AAACAAGAGAGCGGTACGCTGCTGTGTGGAAGGGCAAGACGGAGAGGGCCCAAGCCCA	2272
Qy	2158	GACCTGCTTATCCAGAGGGGGCCACCGCTCTCTGCCCCATGTCACGAGAGTCAAAC	2217
Db	2273	GACCTGCTTATCCAGAGGGGGCCACCGCTCTCTGCCCCATGTCACGAGAGTCAAAC	2332
Qy	2218	AGCTTCCACTTTCATTTGGCTTTCGTTGGAAGAGCAGACTTTTACCACTGTCAAGATTGTC	2277
Db	2333	AGCTTCCACTTTCATTTGGCTTTCGTTGGAAGAGCAGACTTTTACCACTGTCAAGATTGTC	2392
Qy	2278	AATCACTGCTAGCTTTCGCGCTTGGGGGCTCAGGAATGCTTCCCTGCTGCTACCTAT	2337
Db	2393	AATCACTGCTAGCTTTCGCGCTTGGGGGCTCAGGAATGCTTCCCTGCTGCTACCTAT	2452
Qy	2338	ACCAGCTCTGAGAGAGACTTCTCATTTGGGGCTGGAACGNTTTCAGAGTACGAGTTT	2397
Db	2453	ACCAGCTCTGAGAGAGACTTCTCATTTGGGGCTGGAACGNTTTCAGAGTACGAGTTT	2512
Qy	2398	GGGTTACAGTCCCAAGAGTGGATATGGAATGGGGCCCTTTTGGCTCGGTGAGAACGCTCC	2457
Db	2513	GGGTTACAGTCCCAAGAGTGGATATGGAATGGGGCCCTTTTGGCTCGGTGAGAACGCTCC	2572
Qy	2458	ACCTGCTGACCGGCTTCAACACCTCTTCTGACCTGCGCTGAGCGCCCTGACACCA	2517
Db	2573	ACCTGCTGACCGGCTTCAACACCTCTTCTGACCTGCGCTGAGCGCCCTGACACCA	2632
Qy	2518	TCCACCGTTTGGTTTACTGCTGCTCCCCCAAGAGGCGCAATGCTGAGATTGAGATAT	2577
Db	2633	TCCACCGTTTGGTTTACTGCTGCTCCCCCAAGAGGCGCAATGCTGAGATTGAGATAT	2692
Qy	2578	CTAATTTCTTACAGCAACCAACCAACCCAGCGCCGAAACACCAAGTGGACACTGCTCACCA	2637

Db 762 ATGAGTGGGCTCTGTGACCCACCCTTTTGTGTCTGGGTCCGACAGGGGAT 821
Qy 775 ----- 774
Db 822 CAGAGTAAGAGGAACATGGGGATTCAAACACAGCGGCTGTGATGACGCTTTTGCAAGGATGT 881
Qy 775 ----- 774
Db 882 AGTGGAAACCTCGTAGTTGGAGAGGGGTGTAGAGGGGACAGTTCCAGCGTCCCTCCAGC 941
Qy 775 ----- GATGGAAAGCCTATCTCCACGAGTGTATCGTTCTGGGC 814
Db 942 CTGCTCCTTCTGTCAACAGATGGAAGCCTATCTCCACGATGTATCATGTTCTGGGC 1001
Qy 815 GGACCAATCTACTCATCGCCAGCGGAGCTCTGGCACTCTGGAGTCTATGTCCTGGCAG 874
Db 1002 GGACCAATCTACTCATCGCCAGCGGAGCTCTGGCACTCTGGAGTCTATGTCCTGGCAG 1061
Qy 875 CCAACAAAGCCCTCAACGCTGACTTCGCCACTCGCGGCTGTGAGCTCCGAGTGTCTGCTG 934
Db 1062 CCAACAAAGCCCGCACGCGTGTATTTGGCCACTCGCGGCTGTGAGCTCCGAGTGTCTGCTG 1121
Qy 935 CCCAGCCATCTCGAGGCAACCGAGGCGCTCTCGCGACCGCGGCTGAGCACCGCGGCT 994
Db 1122 CCCAGCCATCTCGAGGCGCCCGAGGCGCTCTCGCGACCGCGGCTGAGCACCGCGGCT 1181
Qy 995 TCGTGTGCCGGCGTCCGGGAGCACAGCGCCGCGCTGCACTGGCTGCAAGCGGATCC 1054
Db 1182 TCGTGTGCCGGCGTCCGGGAGCACAGCGCCGCGCTGCACTGGCTGCAAGCGGATCC 1241
Qy 1055 CGTTGCGCAACCAATGGG-CGGGTCAAGGTGAGGCGGCGGTGGCGAGCTTGGTCACT 1113
Db 1242 CGTTGCGCAACCAATGGGCGGTCAGAGTGCAGGGGCTGGCGAGCTTGGTCACT 1301
Qy 1114 CAGATCGGCTTGCAGGACGCTGGCTATACAGTGGGTAGCAGAAAAACAGCGCGGAACT 1173
Db 1302 CAGATCGGCTTGCAGGACGCTGGCTATACAGTGGGTAGCAGAAAAACAGCGCGGAACT 1361
Qy 1174 GCCTGTGCCGCTGGCCCTGGCGGTAGTGTGGCGGGGCTGCCAGCGCCCGACT 1233
Db 1362 GCCTGTGCCGCTGGCCCTGGCGGTAGTGTGGCGGGGCTGCCAGCGCCCGACT 1421
Qy 1234 CGGGTCAAGCAGCAGCCGCTGAGCAGCTCTCTGTGTGGTGGCTGGGAGCGGCTGAG 1293
Db 1422 CGGGTCAAGCAGCAGCCGCTGAGCAGCTCTCTGTGTGGTGGCTGGGAGCGGCTGAG 1481
Qy 1294 TTGCAACAGCAGCAAAATCATTTGGCTTCTCTTCTCACTAACAAAGCAAGGGAGTGGAC 1353
Db 1482 TTGCAACAGCAGCAAAATCATTTGGCTTCTCTTCTCACTAACAAAGCAAGGGAGTGGAC 1541
Qy 1354 AATGTGGAGTACCAAGTTTGCAGTAAACAAATGACACACAGAGCTGCAGGTTCCGGACCTG 1413
Db 1542 AATGTGGAGTACCAAGTTTGCAGTAAACAAATGACACACAGAGCTGCAGGTTCCGGACCTG 1601
Qy 1414 GAAACCAACAGGATATAGTTTCTACGTGTGGCTTACTCCAGCTGGGGGCGAGCCGA 1473
Db 1602 GAAACCAACAGGATATAGTTTCTACGTGTGGCTTACTCCAGCTGGGGGCGAGCCGA 1661
Qy 1474 ACCTCCAGCCAGCCCTGGTGCATACACTGGACAGATGTCCCAGCGCAGCACCCTGAGCTT 1533
Db 1662 ACCTCCAGCCAGCCCTGGTGCATACACTGGACAGATGTCCCAGCGCAGCACCCTGAGCTT 1721
Qy 1534 ACCTTGTCCAGCCCAACCTCCGACATACAGGGTGGATGGCTGCCCTGCCCTCCAGC 1593
Db 1722 ACCTTGTCCAGCCCAACCTCCGACATACAGGGTGGATGGCTGCCCTGCCCTCCAGC 1781
Qy 1594 CTGAGCATGACAGGCTGAGTTTCAAGTACAGATAGAGTACGTTTGGGAAG---GAGAT 1650
Db 1782 CTGAGCATGACAGGCTGAGTTTCAAGTACAGATAGAGTACGTTTGGGAAGGAAAGAT 1841
Qy 1651 CAGGTTTTCTCCACGAGGCTGCTGGAAATGAGACACAACTTACGTTTAAATCTCACTTCAG 1710
Db 1842 CAGGTTTTCTCCACGAGGCTGCTGGAAATGAGACACAACTTACGTTTAAATCTCACTTCAG 1901

Qy 1711 CCAACAAAGTGTACCGAGTCCGATTTCACTGTGGCACTGGCGCTGTATGAGTCCCT 1770
Db 1902 CCAACAAAGTGTACCGAGTCCGATTTCACTGTGGCACTGGCGCTGTATGAGTCCCT 1961
Qy 1771 TCTCAGTGGATGACGACAGGACACCTGGTGTGCAAAACAGAGCCATGTTCCCTTTGCC 1830
Db 1962 TCTCAGTGGATGACGACAGGACACCTGGTGTGCAAAACAGAGCCATGTTCCCTTTGCC 2021
Qy 1831 CTTGCAAGATTTGAAGGTGAGGCAAGATGAGTCCCTGGTGGTGTCACTGGCAGCCGCC 1890
Db 2022 CTTGCAAGATTTGAAGGTGAGGCAAGATGAGTCCCTGGTGGTGTCACTGGCAGCCGCC 2081
Qy 1891 CTTCAACCCCAACAGATCTCTGGATACAACTCTACTCTGGGAGAGTGGGAACAGAGAG 1950
Db 2082 CTTCAACCCCAACAGATCTCTGGATACAACTCTACTCTGGGAGAGTGGGAACAGAGAG 2141
Qy 1951 GAGCAGATGGTGAACCGCCCCCGAGGGGTCTGGAGATCAAGCTTGGAGCTCGGGGCC 2010
Db 2142 GAGCAGATGGTGAACCGCCCCCGAGGGGTCTGGAGATCAAGCTTGGAGCTCGGGGCC 2201
Qy 2011 GTGCGGCTGAAGAGAAAGTGAAGCAGTATGAACCTGACCTGAGTGTCTCTGGCAGGCTG 2070
Db 2202 GTGCGGCTGAAGAGAAAGTGAAGCAGTATGAACCTGAGTGTCTCTGGCAGGCTG 2261
Qy 2071 TACGAGTGAAGTCTGTAGCTTTCAAACAAACAGAGACGCTACGCTGTGTGGAG 2130
Db 2262 TACGAGTGAAGTCTGTAGCTTTCAAACAAACAGAGACGCTACGCTGTGTGGAG 2321
Qy 2131 GGCAAGACGAGAGAGCGGCCCAACGAGACTGTGCTATCAAGAGGGGCGCACCGCTGCT 2190
Db 2322 GGCAAGACGAGAGAGCGGCCCAACGAGACTGTGCTATCAAGAGGGGCGCACCGCTGCT 2381
Qy 2191 CTTGCCATGCTCCACGAGAGTCAACAGCTCCACTTCCATTTGGCTTGGTGGAGAGAG 2250
Db 2382 CTTGCCATGCTCCACGAGAGTCAACAGCTCCACTTCCATTTGGCTTGGTGGAGAGAG 2441
Qy 2251 CCAGACTTTTACCCTGTCAAGATTTGTCAACTACACTGTACGCTTCCGCGCCCTCGGGGCTC 2310
Db 2442 CCAGACTTTTACCCTGTCAAGATTTGTCAACTACACTGTACGCTTCCGCGCCCTCGGGGCTC 2501
Qy 2311 AGGAATGCTTCCCTGTGCTACTATACAGCTCTCGGAGAGACATTTCTCATTTGGGCGC 2370
Db 2502 AGGAATGCTTCCCTGTGCTACTATACAGCTCTCGGAGAGACATTTCTCATTTGGGCGC 2561
Qy 2371 CTGAAACCATTTACCAAGTACGAGTTTGGTTCAGTCCCAAGAGTGGATATGATGGG 2430
Db 2562 CTGAAACCATTTTACCAAGTACGAGTTTGGTTCAGTCCCAAGAGTGGATATGATGGG 2621
Qy 2431 CCTTTGGCTCCGCTGTAGAACGCTCCACCTGCTGACCGGCTTCAACACCTCTTCT 2490
Db 2622 CCTTTGGCTCCGCTGTAGAACGCTCCACCTGCTGACCGGCTTCAACACCTCTTCT 2681
Qy 2491 GACCTGCGCTGAGCCGCCCTGACACCATCCACCTGCTGCTGCTGCTGCTGCTGCTGCT 2550
Db 2682 GACCTGCGCTGAGCCGCCCTGACACCATCCACCTGCTGCTGCTGCTGCTGCTGCTGCT 2741
Qy 2551 GAGCCCAATTTGAGATTTGGAGTATCTAATTTCTTACAGCAACACCAACCCAGGCC 2610
Db 2742 GAGCCCAATTTGAGATTTGGAGTATCTAATTTCTTACAGCAACACCAACCCAGGCC 2801
Qy 2611 GAAACACAGTGGACACTGCTCCACAGAGGGAAACATCTTCAAGATGGAGGTCATGCGC 2670
Db 2802 GAAACACAGTGGACACTGCTCCACAGAGGGAAACATCTTCAAGATGGAGGTCATGCGC 2861
Qy 2671 CTAGAGTGGACACTCGGTATTTCTTCAAGATGGAGGTCATGCGCCTGAGGCTGGG 2730
Db 2862 CTAGAGTGGACACTCGGTATTTCTTCAAGATGGAGGTCATGCGCCTGAGGCTGGG 2921
Qy 2731 CCTTTTCCGCTTGCAGGATGTGATTAATCTCTGCAAGAGACATTTCTCAGACTCTTGGAT 2790
Db 2922 CCTTTTCCGCTTGCAGGATGTGATTAATCTCTGCAAGAGACATTTCTCAGACTCTTGGAT 2981

BASE COUNT		protein"		708 a		1150 g		673 t		1 others	
ORIGIN											
Query Match		74.8%;		Score 2090.8;		DB 6;		Length 3741;			
Best Local Similarity		85.3%;		Pred. No. 0;							
Matches 2394;		Conservative		0;		Mismatch 393;		Indels 18;		Gaps 5;	
QY	1	GGGGAGCTGCCATTGGCC	CAGAGAGCAACTGTCA	AGCTGAGCTGAGCTGTG	ATGAGGAGCCCTG	60					
DB	70	GGGAGCTGCTGTGGCC	CAGAGAGCACTGTG	AGCTGAGCTGTGAGT	TGGGGCCACTG	129					
QY	61	CAAGTGATCCTGGGCC	CTGAGCAGGCTGTGT	GTGAGCTGCACTTT	TGGGGCTACAGCT	120					
DB	130	CAAGTGATCCTGGGCC	CAGCAGGCTGCAGT	GTCTAACTGTAGCCT	TGGGGCTGTGCTC	189					
QY	121	GCTGGGCTTCGAC	ACAGGCTGATGAG	CAAGGATGAGACAC	ACTGTACTAGAGCATG	180					
DB	190	GCTGGACCCCC	CACAGGCTGACCT	TGAGCAAGGATGGG	GCACCTGCTGGAGCAG	249					
QY	181	AACCTGCACCTG	CTACCCCAATGGCT	CCCTGCTGTGCTCT	CACCCCTAGAGCAAGAG	240					
DB	250	CACCTACACCTG	CTGCGCCCAATGGT	TCCCTGTGCTGTCT	CCAGCCTAGCACCCATGGC	309					
QY	241	AGCGATGATGAG	GAAGCTCTTAGGAT	CTGGAAGGCTCACT	GTGAGGCGAGCTATTCTGTCTG	300					
DB	310	AG---TG	AGAGTCACTCT	GTGGGCTCATTT	GAAGGCAACTATTCTGTGCTA	366					
QY	301	GCCACAGCCGCT	AGGAGTGTGGC	CAGCAGGTTGCTGT	GTGTCAGCTTGCACACTC	360					
DB	367	GCCACAGNCC	CTCGAGTGTGGC	CAGCAGACTGCTGT	GTGTCAGCTTGCACACTC	426					
QY	361	GAAGCTTCTCT	CTGACCCCGAGT	CTCCAGATTTGT	GGAGGAGAACGGGACAGCAGCTTT	420					
DB	427	GCAGACTTCT	CTGACCCCGAGT	CTCAGAGCTGT	GGAGGAGAACGGGACAGCTGCTTT	486					
QY	421	GAATGCCACAC	CAAGGCTTCC	CAGCCCTCATTTACT	TTGGGAAAGGACCAAGGTGACC	480					
DB	487	GAGTGCCACAT	TGAAGGCTGCG	CAGCTCCCATCATTTACT	TTGGGAGGAGGACCAAGGTGACA	546					
QY	481	GTGCTTGAGAG	CCCCCGCTCAT	CACTTTCCCAAGT	GTGCTCTCCAGATCTCTAGATGTC	540					
DB	547	TTGCTCTGAG	AGAGCTCGGCTCAT	CGTGTCTTCCCAAGG	GTGCTTCCAGATCTCTGATGTT	606					
QY	541	CAGGACAGTGAT	GCAGGCTTAC	CGCTGGGCGGCA	CCAAATTCACGCCCGCCCAACGATTC	600					
DB	607	CAGGACAGTGAT	GCAGGCTTAC	CGCTGGGCGGCA	CCAAATTCACGCCCGCCCAACGATTC	666					
QY	601	AGCCAGAGGCT	CTCGCTCACTGT	GGCCCTCAGAGGCT	CTTTTGGAGGCTTACAGGGGGCAG	660					
DB	667	AGCCAGAGGCT	CTCGCTCACTGT	GGCCCTCAGAGGCT	CTTTTGGAGGCTTACAGGGGGCAG	726					
QY	661	GATGTGGTCA	TTGTGGAGCC	CCAGAGAACACCA	CGGTAGTGTCTGGAAGAAATGATGTG	720					
DB	727	GACGTGGTCA	TTGTGGAGCC	CCAGAGAACACCA	CGGTAGTGTCTGGAAGAAATGATGTG	786					
QY	721	ATGGAGTGGT	GCTCTGTGACCC	CCACCCCTTTTGT	GTCTCTGGGTCCGACAGGATGGA	780					
DB	787	ATGGAAATGT	GTCGCTTACGCT	GTACCCACCCCTTTTGT	GTCTCTGGGTCCGACAGGATGGA	843					
QY	781	AAGCCTATCT	CCACAGATGTCT	GTGGGCGGACCA	CAATCTACTCTATCGCCAGCGG	840					
DB	844	AAGCCTATCT	CCACAGATGTCT	GTGGGCGGACCA	CAATCTACTCTATTTGCCAACGGG	903					
QY	841	CAGCTCGG	CACTCTGAGT	CTATGTCTGCC	AGCCCAACAGCCCTCTACCGGTGACTTC	900					
DB	904	CAGCCCTGG	CACTCTCGGCT	CTATGTCTGCC	AGCCCAACAGCCCTCTACCGGTGACTTC	963					
QY	901	GCCACTCGG	CTCTGAGTCCG	AG---TGCT	TGTGCTGCCCCCAGCCTCTCGAGGCC	957					
DB	964	GCCACTGCA	CGCCCTGAGCT	CCGCTGTGCTG	CTAGCGGCTCCCGCCATCTACCTAGGCGCCC	1023					
QY	958	GAGGCGCT	CTCTCGGAG	CGGGCGAC	CAACCGCGCTTCTGTGTGCCGGGCTCCGGGGAG	1017					

DB	1024	GAGGCGCT	GTCTCGCGAG	CGCGGCGAG	CACACGCGGCTTCTGTGTGCGCGCTCGGGGAG	1083					
QY	1018	CCAGGCGCG	CGCTGCTCA	CTGCTGCAG	ACCGGATCCCGTTGCGACCAATGGGCGCTC	1077					
DB	1084	CCGCGCG	CAGCGCTG	CGCTGCTGC	CAACCGGGGCGCGCTCGGCGCCACAGGGCGCTC	1143					
QY	1078	AAGGTG	CAGGCGGCTG	CGGCGAGCTT	GGTCACTCACTCAGATCGGCTCTGACGAGACGCTGGC	1137					
DB	1144	AAGGTCC	AGGCGCGCG	CGGTGGCAGCT	TGTCATCACACAGATCGGCTCTGACGAGACGCGCGC	1203					
QY	1138	TACTAC	AGTGTGATG	CAAGAAAC	ACGCGCGGAACTGCTGTGTGCGCTCGCGCCCTGCGG	1197					
DB	1204	TACTAC	AGTGTGCTG	GTGCTGAG	AACAGCGCGGAAATGGCGTGTGCGCTCGCTGTGCTC	1263					
QY	1198	GTAGTGT	GGCGGAGG	GGCTGCCAG	CGCCCGGCTCGGGTCAACAGCCACCGCTGAGC	1257					
DB	1284	GTGTGT	GGCGGAGG	GGCTGCCAG	CGCCCGGCTCGGGTCACTGTCTACCGCTGAGC	1323					
QY	1258	AGCTCT	CTGTGTG	GTGGCGCT	TGGAGCGGCTGTGATTTGCAACGAGCAAAATCAATGGC	1317					
DB	1324	AGCTCG	CTGTGTG	GTGGCGCT	TGGAGCGGCTGTGATTTGCAACGAGCAAAATCAATGGC	1383					
QY	1318	TTCTCT	CTTCACTAC	AAAGGCA	AGGGAAGTGGACAAATGTGGAGTACAGTTTGGCACTA	1377					
DB	1384	TTCTCT	CTTCACTAC	AAAGGCA	AGGGAAGTGGACAAATGTGGAGTACAGTTTGGCACTA	1443					
QY	1378	AACATG	ACACACAG	AGCTG	CGGTTCCGGACCTTGGAAACCCACACAGATTTAGATTTC	1437					
DB	1444	AACAA	CGACACACAG	AACTAC	AGTTTCGGACCTTGGAAACCCACACAGATTTAGATTTC	1503					
QY	1438	TACGTGT	GGGCTTACT	CTCCAG	CTGGGCGCAGCCGAACTCTCAGCCAGCCCTGGTGCAT	1497					
DB	1504	TACGTGT	GGGCTTACT	CTCCAG	CTGGGCGCAGCCGAACTCTCAGCCAGCCCTGGTGCAT	1563					
QY	1498	ACATGGA	CGATGTCCC	CAGCGAC	ACCCAGCTTACTTGTTCAGGCCCAACCCCTCG	1557					
DB	1564	ACATGGA	CGATGTCCC	CAGCGAC	ACCCAGCTTACTTGTTCAGGCCCAACCCCTCG	1623					
QY	1558	GACATCA	GAGTGGCAT	GTGCTGCC	CTCCAGCTTCCAGCTGTGAGCAAGTGGACAGTGTGAG	1617					
DB	1624	GACATCA	GAGTGGCAT	GTGCTGCC	CTCCAGCTTCCAGCTGTGAGCAAGTGGACAGTGTGAG	1683					
QY	1618	TACAAG	ATAGAGTAC	CGTTTGGG	GAAGAAAGATCAGGTTTCTCCACCGAGGTGCTGGA	1677					
DB	1684	TACAAG	ATAGAGTAC	CGTTTGGG	GAAGAAAGATCAGGTTTCTCCACCGAGGTGCTGGA	1743					
QY	1678	AATGAG	ACACAACTT	TAACCTCA	CTTCAGCCAAACAAAGTGTACCGAGTCCGAT	1737					
DB	1744	AATGAG	ACACACAG	CTTATG	CTGAACTCGCTTCAAGCCAAACAGGTGTATCGAGTACGGATT	1803					
QY	1738	TCAGCT	GGCCTG	GGCTGGCT	TATGAGTTCCTTCTCAGTGTGATCGACACAGGACCT	1797					
DB	1804	TCGGCT	GTGTACAG	AGCGGCTT	CGGGGCGCCCTCCAGTGTGATCGACACAGGACCT	1863					
QY	1798	GGTGTG	CAACACAG	AGCATGTT	CCCTTTGGCCCTCTGACAGATTTCAAGGTGAGGGCAAG	1857					
DB	1864	AGTATG	CAACACAG	AGCATGTT	CCCTTTGGCCCTCTGACAGATTTCAAGGTGAGGGCAAG	1923					
QY	1858	ATGAGT	CCCTGTG	GTGTCA	TGTCAGCGCGCCCTCTACCCCAACAGATCTCTGATAC	1917					
DB	1924	ATGAGT	CCCTGTG	GTGTCA	TGTCAGCGCGCCCTCTACCCCAACAGATCTCTGATAC	1983					
QY	1918	AAACTCT	ACTTGGG	GAGGTGGAA	CAGAGGAGGAGGAGTGTGACCGCCCGCCCGG	1977					
DB	1984	AAACTCT	ACTTGGG	GAGGTGGAA	CAGAGGAGGAGGAGTGTGACCGCCCGCCCGG	2043					
QY	1978	GGTGTG	GAGATCA	AGCTTGG	GACCTCGGCGCTGTGAGTGTGAGGAGGAGGAGTGTGAGGAG	2037					
DB	2044	GGCGT	GAGAGC	AGGCTTGG	AGTGTGGGCTGTGAGGAGGAGGAGTGTGAGGAG	2103					
QY	2038	TATGAA	CTGAC	CCAGTTAGT	CTCCCTGGCAGCGGCTACGAGGTGAAGCTCGTAGCTTTCAAC	2097					

Db 2104 TATGAGCTGACCCAGCTAGTCCCTGGCGGGCTGTACAGGTGAAGCTGTGGCTTTTCAC 2163
Qy 2098 AAACAAGGAGCGCTACGCTGTGTGTGGAGGGCAAGACGGAGAGGGCCCAACCCCA 2157
Db 2164 AAACATGAGGATGCTATGACGAGCTGTGTGAAGGGCAAGACGGAGAGGGCCCAACCA 2223
Qy 2158 GACCTGCTATCCAGAGGGGGCCACCGCTGCTCTCTGCTCCCTGCTCCAGCAGAGTCAAC 2217
Db 2224 GACATGCTATCCAGAGGGGGACACCCCTGCTCCAGCCCACTGCTGCGGAATCAAC 2283
Qy 2218 AGCTCCACTCTCCATTTGCTGTGGGTGAAGAGGACAGCTTTTACCACTGTCAAGATTGTC 2277
Db 2284 AGCTCCACATCCATCTGGCTTGGGTGAAGAGGACAGTTTCCACACAGTCAAGATTGTC 2343
Qy 2278 AACTPACCTGPAACCTTTGGGCCCTGGGGGCTCAGGAATGCTTCCCTGGTCACTACTAT 2337
Db 2344 AACTPACCTGTCGCTTTGAGCCCTGGGGGCTCAGGAATGCTTCCCTGGTCACTACTAT 2402
Qy 2338 ACCAGCTCTGGAGAGACATCTCTATTTGGGGGCTGAAACATTTTACCAAGTACGATTT 2397
Db 2403 --CAGTTCTGGAGAGACATCTCTATTTGGGGGCTTGAAGCCATTTCAACAAATACGATTT 2460
Qy 2398 GCGGTACAGTCCACGAGGTGATGATGATGGGCTTTGGCTCGCTGCTAGAACGCTCC 2457
Db 2461 GCACTGAGTCTACGGGCTGGAATGATGGGCTTTGGCTCTGTGGTGAAGCGCTCC 2520
Qy 2458 ACCCTGCTGACCGGCTTTCAACACCTCTTTGACCTGGCCCTGAGCCCTGACACCA 2517
Db 2521 ACCCTGCTGACCGTCTCCACACACCCCTCCGACCTCGGACTGAGCCCTGACACCG 2580
Qy 2518 TCCAGCTTGGTTTACACTGTTGCTCCCAACGAGCCCAATGTTGATGTTGGAGTAT 2577
Db 2581 TCCAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
Qy 2578 CTAATCTCTACAGCAAC 2636
Db 2641 CTGATCTGTACAGCAAC 2700
Qy 2637 -----AGAGGGAACATCTTTCAGTGACAGAGTCCATGGCTTAGAGAGTGCACCTCGGTAT 2691
Db 2701 CAGGCTGAGGGAACATCTTTCAGTGCTGAGTCCATGGCTTAGAGAGTGCACCTCGGTAT 2760
Qy 2692 TTCTTCAAGTGGAGCGCCGACAGAGTGGGGCTGGGCCCTTTTCCCGCTTGCAGGAT 2751
Db 2761 TTCTTCAAGTGGGGCGGACAGAGTGGGACCTGGGCTTTCTCCCGCTTGCAGGAT 2820
Qy 2752 GTGATCTCTGACAGAGACATCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCA 2796
Db 2821 GTGATCAGCTCCAGGAGAGCTGTGACAGTCTGACAGTCTGACAGTCTGACAGTCTG 2865

RESULT 10
LOCUS AB046848 5549 bp mRNA linear PRI 22-FEB-2001
DEFINITION Homo sapiens mRNA for KIAA1628 protein, partial cds.
ACCESSION AB046848
VERSION AB046848.1 GI:10047332
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Nagase, T., Kikuno, R., Nakayama, M., Hirose, M. and Ohara, O.
TITLE Prediction of the coding sequences of unidentified human genes.
XVII. The complete sequences of 100 new cDNA clones from brain
which code for large proteins in vitro
JOURNAL DNA Res. 7 (4), 273-281 (2000)
MEDLINE 20450683
PUBMED 10997877
REFERENCE 2 (bases 1 to 5549)
AUTHORS Ohara, O., Nagase, T. and Kikuno, R.
TITLE Direct Submission

JOURNAL Submitted (03-AUG-2000) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdna@kazusa.or.jp,
URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,
Fax:81-438-52-3914)
FEATURES
source Location/Qualifiers
1. .5549
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="fh14247"
/notes="vector:pBluescriptII SK plus"
gene 1. .5549
/genes="KIAA1628"
CDS
1. .3023
/genes="KIAA1628"
/note="Start codon is not identified. fh14247 cDNA clone
for KIAA1628 has a 40-bp different sequence replacing a
108-bp sequence at the position between 737 and 846, and
an 88-bp insertion after the position 983 of the sequence
of KIAA1628, respectively."
/codon_start=1
/product="KIAA1628 protein"
/protein_id="BAB13454.1"
/db_xref="GI:10047333"
/translation="MGRKGRSALSSRLSVNQEVKVIIGRGAREGGGQAWGTGGG
GDLSCGIPRSAPLAIPAIQAPALSRSTARFVCRASGPPALRWLHNG
APLPNGRVKVOGGSLVITQIGLDAGYOCVAENSAGMACASLAVVREGLPS
APRTVATPLSSAVLWAMERPMHSEBQIIGSLHYQKAGMDNVEYQFVANDTTEL
QVRDLNPWDYFYVAYSQLGASRTSTPALVHLLDDVPSAQLSPNSFSDIRVA
WLPPLPSLNGQVYKIKYIGLIGKQDIIFSTEVRNETMLNSLQPNKRVIRISG
TAAGFGAPSGMHRTTPSHNSHPVFPAPBLKVAQKESLVSMQPPPHPTQISYK
LYWREVGAREAGDRLPGCRGDQAWDVGVPRLKKVKQVYELTQLVGLYKLVAF
NKHEDGYAAVWKGTEKAPADMPIDRGPPLPAHVHASENSSTSLWLSKKDPDTTV
KLNVTVRSPWGLNASLVYTTSSGEDIILGGLKPTPKYFPAVSHGVHMDGPPGS
VVERSLPDRSPDPSLPLTSTVRLHWCPTTPEGEIVYDILYSSNHTPEH
QMTLLTQGNIFSAEVHGLSDTRYFFKMGARTVEGPPFRLQDVTIQLKLSLSD
MSVTHIIVGCLGLLCLLACAGLRSPHRESLPLSLSTATPGNPALYSRLALGP
SPAAHLSLSLPHPPQDPPSPDVDEARVHSLMGVSGRSHSKKISWAQPSG
LWPAHLSLSLPHPPQDPPSPDVDEARVHSLMGVSGRSHSKKISWAQPSG
NOVEARVIVHSDPSASNGNDLHLQDLEPDEPLPPEAPDLISGVGDPGGAALWDLRL
GCGLAAPGPDRLTCLLPEAASASCYPDLPQPEVLSETPGDSCQLKSPCLGASPLP
RSPVSSSA"

BASE COUNT 1157 a 1546 c 1496 g 1350 t
ORIGIN

Query Match 53.3%; Score 1491.4; DB 9; Length 5549;
Best Local Similarity 87.4%; Pred. No. 3, 3e-298;
Matches 1633; Conservative 0; Mismatches 236; Indels 0; Gaps 0;

Qy 928 CTGTGCTGCCCGCAGCATCTCGAGGCAACCGAGGCGCTCTCGCGGAGCGGGCGCAGCAC 987
Db 264 CTGAGCGCTCCCGCCATCACTCAGGCGCCCGAGCGCTGTCCGCGAGCGGGGAGACA 323
Qy 988 GCGCGCTTCTGTGTCGGCGGTCTCGGGGAGCCACGCGCCGCGCTGACCTGCTGCTGAC 1047
Db 324 GCGCGCTTCTGTGTCGGCGGTCTCGGGGAGCGCGCGCTGCGCTGCTGCTGCAAC 383
Qy 1048 GCGATCCGCTTGGACCAATGGCGCGCTCAAGGTGACGGCGGTGCGCGCAGCTTGGTC 1107
Db 384 GGGCGCGCTTGGCGCCAAACGCGCGCTCAAGGTGACGGCGGTGCGCGCAGCTTGGTC 443
Qy 1108 ATCAGCTCAGATCGGCTGACAGGAGCTGTGCTACTACAGTGGCTAGCAGAAACAGCGG 1167
Db 444 ATCAGCTCAGATCGGCTGACAGGAGCTGTGCTACTACAGTGGCTGCTGAGAAACAGCGG 503
Qy 1168 GGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1227
Db 504 GGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 563
Qy 1228 CCGACTCGGGTCAACGCCAGCGCGCTGAGCAGCTCTCTGTGCTGTGCTGTGCTGTGCTG 1287
Db 564 CCGACGCGGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 623

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Deng, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackeon, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C., Labow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuewa, L., Loulesged, H., Lozada, R., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Naif, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norrie, S., Nwankwelu, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., S.J., Sanders, M., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C., D., Smajs, D., Sneed, A., Sodergren, E., Song, X., Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, S., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstein, G. and Gibbs, R. A.

Direct Submission
Unpublished
2 (bases 1 to 224227)
Worley, K. C.

Submitted (12-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 224227)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 13, 2003 this sequence version replaced gi:22857495.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine

Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: CH230-20218
Center clone name: GIBA
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 211258 bases at least Q40
Consensus quality: 214044 bases at least Q30
Consensus quality: 215992 bases at least Q20
Estimated insert size: 220399; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 221791: contig of 221791 bp in length
* 221792 222909: contig of 1018 bp in length
* 222910 223009: gap of unknown length
* 223010 224227: contig of 1218 bp in length.
FEATURES
source
1..224227
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-20218"
1..1752
/note="wgs end extension
clone end:Sp6"
3617..3835
/note="clone boundary
clone end:Sp6"
site:EcoRI
end:sequence:BH334357"
220120..221791
/note="wgs contig"
misc_feature
58284 a 52300 c 50849 g 55959 t 6835 others
BASE COUNT
ORIGIN
Query Match 13.5%; Score 376.8; DB 2; Length 224227;
Best Local Similarity 88.7%; Pred. No. 7, 5e-68;
Matches 408; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 919 CTCGAGTGTCTGTCGCCCGACGCAATCTGCGAGGACCCGAGCGCTCTCGCGGACGCG 978
Db 109894 CTCCTGCTCTCAGCTGCTCCCGACGCAATCTGCGAGGCGCCGAGCGCTCTCGCGGACGCG 109835
QY 979 GCCAGACCGCGCGCTTCTGTCGCGGGGTCGCGGGGACCCAGCGCGCTGCTGCTGCTGCTG 1038
Db 109834 GCCAGACTGCGCGCTTCTGTCGCGCGCTTCTGCGGGGACCCAGCGCTGCTGCTGCTGCTG 109775
QY 1039 CTGCGACGACGGGATCCGCTTTCGCGACCCAAATGCGCGCGCTCAAGCTGCGAGGCGCTGCGCG 1098
Db 109774 CTGCGACGACGGGCGCGCTTTCGCGACCCAAATGCGCGCGCTCAAGCTGCGAGGCGCTGCGCG 109715
QY 1099 AGCTTGGTCACTACATGATCGGCTGCGAGGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1158
Db 109714 AGCTTGGTCACTACATGATCGGCTGCGAGGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 109655
QY 1159 AACAGGCGGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1218
Db 109654 AACAGGCGGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 109595
QY 1219 CCCAGCGCCCGACTCGGCTGCGGCTCACAGCCCGCGCTGAGCAGCTCCTCTGCTGCTGCTGCTG 1278

Db 109594 CCCAGTCCCGGCTCGGCTCAAGCAACGCGCTGAGCAGCTCGCTGTGTAGTGGCC 109535
QY 1279 TGGGAGCGGCTGAGTTCACAGCAGCAATCATTTGGCTTCTCTTCACTACCAAAAG 1338
Db 109534 TGGGAGCGGCTGAGTTCACAGCAGCAATCATTTGGCTTCTCTTCACTACCAAAAG 109475
QY 1339 GCAAGGGAGTGGACAATGTGGAGTACCAGTTTGCAGTAA 1378
Db 109474 GCAAGGGAGTGGCTACTTGTACCAACGCGGACCCAGTTA 109435

RESULT 14
AC068507/c
LOCUS
DEFINITION Homo sapiens chromosome 15 clone RP11-349H15 map 15, WORKING DRAFT
ACCESSION AC068507.2 GI:8389522
VERSION AC068507.2
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Birren,B., Linton,L., Nusbaum,C. and Lander,B.
Homo sapiens chromosome 15, clone RP11-349H15
Unpublished
2 (bases 1 to 176282)
Birren,B., Linton,L., Nusbaum,C., Lander,B., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bada,P.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczyk,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheters,R.,
Meidrin,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Teefaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (03-MAY-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 9, 2000 this sequence version replaced gi:7684421.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L10275
Center clone name: 349_H15
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 156826 bases at least Q40
Consensus quality: 166608 bases at least Q30
Consensus quality: 170674 bases at least Q20
Insert size: 173682; sum-of-contigs

Quality coverage: 3.6 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 27 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1477: contig of 1477 bp in length
* 1478 1577: gap of 100 bp
* 1578 2785: contig of 1208 bp in length
* 2786 2885: gap of 100 bp
* 2886 4859: contig of 1974 bp in length
* 4860 4959: gap of 100 bp
* 4960 6416: contig of 1457 bp in length
* 6417 6516: gap of 100 bp
* 6517 8739: contig of 2223 bp in length
* 8740 8839: gap of 100 bp
* 8840 10868: contig of 2029 bp in length
* 10869 10968: gap of 100 bp
* 10969 13737: contig of 2769 bp in length
* 13738 13837: gap of 100 bp
* 13838 15329: contig of 1492 bp in length
* 15330 15429: gap of 100 bp
* 15430 17938: contig of 2509 bp in length
* 17939 18038: gap of 100 bp
* 18039 19881: contig of 1943 bp in length
* 19882 20081: gap of 100 bp
* 20082 23587: contig of 3506 bp in length
* 23588 23687: gap of 100 bp
* 23688 27447: contig of 3760 bp in length
* 27448 27547: gap of 100 bp
* 27548 30161: contig of 2614 bp in length
* 30162 30261: gap of 100 bp
* 30262 34737: contig of 4476 bp in length
* 34738 34837: gap of 100 bp
* 34838 40101: contig of 5264 bp in length
* 40102 40201: gap of 100 bp
* 40202 44039: contig of 3838 bp in length
* 44040 44139: gap of 100 bp
* 44140 49550: contig of 5811 bp in length
* 49551 50050: gap of 100 bp
* 50051 55555: contig of 5505 bp in length
* 55556 55655: gap of 100 bp
* 55656 60496: contig of 4841 bp in length
* 60497 60597: gap of 100 bp
* 60597 68544: contig of 7948 bp in length
* 68545 68644: gap of 100 bp
* 68645 76940: contig of 8296 bp in length
* 76941 77040: gap of 100 bp
* 77041 85480: contig of 8440 bp in length
* 85481 85580: gap of 100 bp
* 85581 93905: contig of 8325 bp in length
* 93906 94005: gap of 100 bp
* 94006 108063: contig of 14058 bp in length
* 108064 108163: gap of 100 bp
* 108164 123496: contig of 15333 bp in length
* 123497 123597: gap of 100 bp
* 123597 139965: contig of 16369 bp in length
* 139966 140065: gap of 100 bp
* 140066 176282: contig of 36217 bp in length.

FEATURES
source
1..176282
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-349H15"
/clone_lib="RPC1-11 Human Male BAC"
misc_feature 1..1477


```
repeat_region 16409..16481
/rpt family="MERSA"
repeat_region 16482..16779
/rpt family="AluSq"
repeat_region 16780..16874
/rpt family="MERSA"
repeat_region complement(17247..17543)

Query Match      11.8%; Score 331.2; DB 9; Length 198295;
Best Local Similarity 80.6%; Pred. No. 2.1e-58; Indels 0; Gaps 0;
Matches 387; Conservative 0; Mismatches 93;

QY 867 CTGCCGAGCAACAAGCCCTCACGCGTCACTTCGCCACTGCGCGTGTGAGCTCCGAGT 926
Db 89528 CTGGGGGGGGGAGAGACTTATTGTCGGGATCCGCGCTCAGCGTCGCCCTCGCGCC 89587

QY 927 GCTTGTGCCCCAGCCATCTGCGAGGCAACCGAGGCGCTCTCGCGGACGCGGGCCAGCAC 986
Db 89588 CCTAGCGGCTCCCGCATCACTCAGCGGCCGAGGCGCTGTGCGGACGCGGGCGAGCAC 89647

QY 987 CGCGCGCTTCGTGTGCGGGCGTCCGCGGAGCCACGCGCGCGCTGCACTGGCTGCACGA 1046
Db 89648 AGCGCGCTTCGTGTGCGGGCGTCCGCGGAGCCGCGCGCTGCGCTGGCTGCACAA 89707

QY 1047 CGGGATCCCGTTGCGCAACCAATGGCGCGCTCAAGGTGCAGGCGGTCGCGCGAGCTTGGT 1106
Db 89708 CGGGGCGCGCTGCGGCGCCNA CGGGCGCGTCAAGGTCCAGGGCGGCGTGGCAGCTGGT 89767

QY 1107 CATCACTCAGATCGGCTCGAGGACGCTGGCTACTACCAAGTGGTAGCAGAAACAGCGC 1166
Db 89768 CATCACACAGATCGGCTCGAGGACGCGGCTACTACCAAGTGGTAGCAGAAACAGCGC 89827

QY 1167 GGGNACTGCTGTGCGCTGCGGCCCTGGGGTAGTGGTGGCGAGGGGCTGCCAGCGC 1226
Db 89828 GGGNACTGCGGTGCGCTGCGGCCCTGGGGTAGTGGTGGCGAGGGGCTGCCAGCGC 89887

QY 1227 CCGCACTCGGGTCACAGCCACGCGCGTGAAGCAGCTCTCTGTGCTGGTGGCTGGGAGCG 1286
Db 89888 CCCCACGCGGGTCACCTGCTACGCCACTGAGCAGCTCCGCTGTGTGGTGGCTGGGAGCG 89947

QY 1287 GCTGAGTTGCACAGCGAGCAATCATTTGGCTTCTCTTCACTACCAAAAGGCAAGGGG 1346
Db 89948 GCGCGAGATGCACAGCGAGCAGATCATCGGCTTCTCTCTCACTACCAAGGCAAGGGG 90007
```

Search completed: October 15, 2003, 04:35:54
Job time : 10466.9 secs